

Regularisierte nichtnegative Matrixfaktorisierungen
und ihre
Anwendungen in der Spektroskopie

Dissertation
zur
Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
der Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Rostock

vorgelegt von
Mathias Sawall, geb. am 31.05.1982 in Wismar

Rostock, 04.02.2011

Gutachter: Prof. Dr. K. Neymeyr, Universität Rostock,
Prof. Dr. R. Ludwig, Universität Rostock

Tag der Verteidigung: 28.04.2011

Inhaltsverzeichnis

1	Einleitung	1
2	Das Faktorisierungsproblem	5
2.1	Faktordarstellungen und Faktorisierungsaufgaben	5
2.2	Die Faktorisierungsaufgabe als schlecht gestelltes Problem	8
2.3	Spektroskopie, Chemometrie und approximative Faktorisierungen	10
3	Die kontinuierliche Faktorisierungsaufgabe	13
3.1	Zusammenhänge zwischen den Faktorisierungen und ihre Einteilung in Äquivalenzklassen	13
3.2	Nichteindeutigkeit der Faktorisierung	16
3.3	Einschränkungen der Menge aller Faktorisierungen	21
4	Die nichtnegative Matrixfaktorisierung	31
4.1	Grundlagen zur nichtnegativen Matrixfaktorisierung	31
4.2	Niedrigrangapproximation einer Matrix	32
4.3	Matrixfaktorisierung mittels Transformation der Singulärvektoren	36
4.4	Nichteindeutigkeit der nichtnegativen Matrixfaktorisierung	38
4.5	Lösungseinschließungen unter zusätzlichen Restriktionen	42
5	Klassische Zugänge zur Matrixfaktorisierung	45
5.1	Hauptkomponenten- und Unabhängigkeitsanalyse	45
5.2	Klassische Algorithmen zur nichtnegativen Matrixfaktorisierung	49
5.3	Selbstmodellierende Zugänge in der Chemometrie	53
5.4	Kritische Zusammenfassung der vorgestellten Lösungsansätze	58
6	Regularisierte nichtnegative Niedrigrangapproximationen und ihre Implementierung in der Methode Pure Component Decomposition	61
6.1	Die Straf- und Regularisierungsfunktionen	63
6.2	Dynamik eines Matrixfaktors und kinetische Analyse	69
6.3	Methoden der nichtlinearen Ausgleichsrechnung	72
6.4	Anwendung von Multi-Level-Methoden	78
6.5	Strategie der sukzessiven Faktorisierung	83
6.6	Gewichtung der Regularisierungsfunktionen	86
6.7	Programmstruktur und Implementierung	88
6.8	Kritische Bewertung des Algorithmus	89
7	Kinetische Modellanpassung	91
7.1	Die Anwendung der regularisierten Niedrigrangapproximation	92
7.2	Eine Folge von Anfangswertproblemen	92
7.3	Die Implementierung	95
8	Ausblick	97

Anhang

A	Optische Spektroskopie	101
A.1	Die FT-IR Spektroskopie	101
A.2	Die UV/Vis Spektroskopie	103
B	Numerische Ergebnisse der regularisierten Niedrigrangapproximation	105
B.1	Ein Modellproblem	105
B.2	Die Hydroformylierung von 3,3-Dimethyl-1-buten	109
B.3	Die Bildung von Hafnazyklopentan	112
B.4	Die Clusterbildung aus $\text{Rh}(\text{acac})(\text{CO})_2$	116
B.5	Analyse der algorithmischen Ideen	119
B.6	Numerische Ergebnisse für eine Systemerweiterung	124
B.7	Kritisches Fazit	129
C	Numerische Ergebnisse zur Systemanpassung	131
C.1	Die Anpassung eines kinetischen Modells zum System aus Anhang B.2	131
C.2	Die Anpassung eines kinetischen Modells zum System aus Anhang B.3	133
D	Ergänzungen und Diagramme	135

1 Einleitung

Der Gegenstand dieser Arbeit ist die numerische Berechnung regularisierter nichtnegativer Matrixfaktorisierungen zu vorgegebenen Matrizen spektroskopischer Messdaten.

Diese Matrizen spektroskopischer Daten und das resultierende Faktorisierungsproblem stehen in engem Bezug zum Gesetz von Lambert und Beer. Bei der Analyse eines chemischen oder physikalischen Systems mittels absorptionsspektroskopischer Techniken bestimmen die Konzentration und die Absorptivität einer Spezies durch ihr Produkt die wellenlängenabhängige Absorption elektromagnetischer Strahlung. Für ein Gemisch mehrerer Komponenten überlagern sich die Absorptionsbeiträge im Idealfall additiv. Untersucht wird ein dynamisches System, in dem sich die Konzentrationen der beteiligten Substanzen zeitlich ändern.

Ausgehend von diesem Zusammenhang wird folgendes Problem untersucht. Gegeben ist einzig eine Matrix von Messdaten zu einem Zeit- und Frequenzraster. Es sind von dem zugrunde liegenden System weder die Anzahl der beteiligten Reinsubstanzen noch deren Spektren bekannt. Auf diese unbekannten Größen soll allein aus der Matrix der Messwerte geschlossen werden.

Dazu ist die Matrix mittels mathematischer Methoden möglichst exakt und aufwandseffizient derart zu faktorisieren, dass die so gewonnene Faktorisierung für das chemische Untersuchungssystem sinnvolle Strukturaussagen zulässt. Diese Strukturaussagen betreffen sowohl die Zahl der teilhabenden Substanzen und deren Reinkomponentenspektren als auch die zugehörigen (zeitlichen) Konzentrationsprofile. Aus physikalischen Gründen sind bei einer Zerlegung nur nichtnegative Faktoren sinnvoll.

In diversen Theorien von Mechanik, Thermodynamik bis hin zur Ökonomie hat es sich als erfolgreich erwiesen, zunächst proportionale, also homogene und additive, Zusammenhänge zwischen abhängigen und unabhängigen Größen zu unterstellen. Entsprechend weit fortgeschritten ist die Theorie linearer Abbildungen und die Numerik zur Lösung von

$$Ax = b \tag{1.1}$$

nach x , sei es durch direkte oder iterative Verfahren. Auch wenn sich im Zuge weiterer Erkenntnisse erweist, dass bei globaler Betrachtung (1.1) nicht exakt gilt, so ist die Beherrschung der linearen Theorie stets der Schlüssel zur Lösung nichtlinearer Erweiterungen. Gleichungen vom Typ (1.1) treten beispielsweise bei der Diskretisierung von Differentialgleichungen auf, [1, 30, 81]; die Matrix A ist wohlbekannt, wenngleich in manchen Anwendungen aufwendig zu berechnen. Ist die Matrix A allerdings nicht bekannt, stellt sich das Problem diese zu identifizieren. Offenbar kann dies nicht gelingen, sofern nur eine rechte Seite b bekannt ist, selbst falls das dazu entsprechende x zugänglich ist. Die Gleichung (1.1) wird auf mehrere rechte Seiten, vereint in der Matrix B , zu $AX = B$ erweitert. Zusätzlich sei an dieser Stelle die Notation zu

$$XY = A \tag{1.2}$$

geändert, um dem Kontext der gegebenen Daten A und den gesuchten Größen X, Y möglichst nahe zu kommen.

Es ist anzumerken, dass, sofern ein Faktor bekannt ist, sei dies beispielsweise X , unter entsprechenden Rangbedingungen an X das Identifikationsproblem für Y eine gut beherrschte Standardaufgabe ist, [5, 55]. Allerdings ist bei der vorliegenden Aufgabe weder X noch Y bekannt.

Vielmehr stellt sich aus dem Anwendungskontext (spektroskopische Daten) heraus noch eine weitere Frage. Die Dimensionen von A sind wohlbekannt, die Spaltenanzahl von X und die Zeilenanzahl von Y jedoch nicht. Liegen spektroskopische Daten A einer chemischen Reaktion vor, so steht die Anzahl der Spalten in X und der Zeilen in Y , folgend s genannt, für die Anzahl der an der Reaktion beteiligten unabhängigen Komponenten. In diesem Zusammenhang bedeutet *unabhängig*, dass sowohl die vektoriellen Einträge der Konzentrationsprofile als auch die der Reinsubstanzspektren jeweils untereinander linear unabhängig sind. Unter entsprechenden Bedingungen an die Dimensionen der Messreihe ist für den Ansatz (1.2) die Anzahl s gleich dem Rang von A .

Die Annahme (1.2) ist oft nur ein idealisiertes Modell; daher sei realitätsnah der multivariate Zusammenhang

$$A = XY + E \tag{1.3}$$

angesetzt. Dabei ist $A \in \mathbb{R}^{k \times n}$ gegeben und es sind die Matrizen $X \in \mathbb{R}^{k \times s}$ und $Y \in \mathbb{R}^{s \times n}$ gesucht, so dass das Residuum E gegenüber dem Produkt XY in einem geeigneten Sinne klein ist. Ferner sollen die Faktoren X und Y von vollem Rang s sein und dieser aber nicht größer als der von A .

Dies ist eine Aufgabenstellung der multivariaten Analysis, deren Arbeitsgebiet die Analyse einer Menge von mehrdimensionalen Daten mittels statistischer Verfahren miteinschließt. Klassische Lehrbücher der multivariaten Analysis sind [9, 40, 42, 43, 52].

Anwendungen des in (1.3) beschriebenen Zusammenhangs finden sich etwa in der Datenkompression (Bildverarbeitung), der Detektion struktureller Zusammenhänge (Klassifizierung von Daten, Abhängigkeitsanalysen) und deren Kombination (Schrifterkennung, allgemein Mustererkennung). In dieser Arbeit liegt der Fokus auf der Analyse und Berechnung multivariater Zerlegungen vom Typ (1.3) mit Anwendungsbezug zur Spektroskopie, weshalb im weiteren Verlauf nur noch nichtnegative Faktoren von Interesse sind.

Zunächst gilt es die idealisierte Aufgabenstellung aus (1.2) in Bezug auf Existenz und Eindeutigkeit einer Lösung mit nichtnegativen Faktoren X, Y vollen Ranges zu untersuchen. Übersichtsarbeiten zur nichtnegativen Matrixfaktorisierung (mitunter ohne Forderung des vollen Ranges der Faktoren) sind [8, 10, 13, 67, 83, 84]. Die Existenz einer Zerlegung ist bei der in dieser Arbeit untersuchten Anwendung gesichert, denn A ist nach der Modellgrundlage gerade das Produkt zweier Faktoren vollen Ranges. Die Frage nach der Eindeutigkeit der Faktorisierung bleibt und mit ihr, ob und wie sich, für den Normalfall einer nichteindeutigen Zerlegung, Mengen möglicher Lösungen determinieren lassen.

Zur näherungsweisen Zerlegung einer nichtnegativen Matrix in nichtnegative Faktoren sind eine Reihe iterativer Methoden bekannt. Diese klassischen Zugänge sind grundsätzlich in drei Klassen zu unterteilen: multiplikative Korrekturformeln, [3, 58–60, 75], Gradientenverfahren, [3, 23, 58, 60, 75] und alternierende kleinste Quadrate Methoden, [3, 47, 60].

Diese Iterationsmethoden erweisen sich auch als partiell erfolgreich, ihr Ergebnis hängt jedoch kritisch von der Wahl der Startnäherung ab. Darüber hinaus liefern die Algorithmen bei der Anwendung auf spektroskopische Daten nur sehr bedingt nutzbare Faktorisierungen.

Speziell auf nichtnegative Matrixfaktorisierungen mit Anwendungsbezug zur Spektroskopie zugeschnittene Algorithmen sind beispielsweise in [63, 86], algorithmische Ideen in [6, 11, 68, 88] zu finden. Wie sich bei einer detaillierten Untersuchung der Methoden aus [63, 86] später zeigen wird, ist deren Erfolg jedoch stark von der Struktur der Ausgangsfaktoren X und Y und damit der Matrix A abhängig; nur ein dünn besetzter Faktor Y (beziehungsweise X bei Transposition der gesamten Problemstellung) begünstigt eine physikalisch-chemisch sinnvolle Zerlegung. Weiter werden nur lokale Niedrigrangapproximation berechnet, in der Form, dass die Faktoren stück-

weise nacheinander als Paare $(X_{:,i}, Y_{i,:})$ extrahiert werden. Dies steht im klaren Widerspruch zur Komplexität der Aufgabenstellung; ohne jede Grundlage wird so (für den idealisierten Fall) unterstellt, dass für die verbleibende Matrix $A - X_{:,i}Y_{i,:}$ exakte Zerlegungen in nichtnegative Faktoren vollen Ranges (nun $s - 1$) existieren und, dass falls es welche gibt, darunter zudem noch mindestens eine physikalisch-chemisch sinnvolle ist.

Der in dieser Arbeit neu entwickelte Algorithmus zur Berechnung einer Faktorisierung basiert auf einer Idee von Lawton und Sylvestre aus [56]. Kernstücke der Methode sind die Singulärwertzerlegung, [18, 22], von A und ein Fehlerfunktional, welches unter anderem die Einträge der Fehlermatrix E , aber auch den Charakter der Faktoren X und Y bewertet. Ein solcher Kompromiss zwischen Genauigkeit – soll doch eigentlich (1.2) und nicht (1.3) gelöst werden – und Einhaltung von vernünftigen Grenzen bei Werten und Variabilität der Faktoren X und Y ist wiederum typisch für schlecht gestellte Probleme, [19, 36, 61], und stellt eine Regularisierung der Aufgabe dar.

Bei dem hier verfolgten Zugang liefert eine abgeschnittene Singulärwertzerlegung die Grundlage geeigneter, niedrigdimensionaler Untervektorräume für die gesuchten Faktoren. Im Normalfall sind diese Faktoren keinesfalls eindeutig, abgesehen von Sortierung und Normierung existieren weitere Freiheitsgrade, deren Parameter im Anschluss mittels Minimierung einer nichtlinearen Fehlerfunktion optimal bestimmt werden.

Angesichts der Schwierigkeiten bei der Lösung von (1.2) beziehungsweise (1.3), die dank des in der Arbeit neu entwickelten Algorithmus zwar reduziert, aber wegen der Schlechtgestelltheit nicht eliminiert werden können, ist es hilfreich, die Aufgabe durch die Einbeziehung weiterer Informationen zu modifizieren. Solche Informationen können zum Beispiel partielle Kenntnisse über beteiligte Substanzen oder über deren Kinetik beinhalten.

Das multivariate Modell (1.3) ist Grundlage vieler weiterer Anwendungen und besitzt zahlreiche Verallgemeinerungen, etwa auf drei oder mehr Faktoren (siehe dazu beispielsweise [48, 51, 66, 77]). Bei dem hier vorliegenden Anwendungsfeld und dem dazu entwickelten Lösungsverfahren hat sich jedoch gezeigt, dass die Feineinstellung der Regularisierungen entscheidend für die Qualität der gewonnenen Faktorisierung ist. Der Algorithmus bietet zahlreiche Optionen und Einstellungen; dies ist einerseits gut für ein interaktives Programm, verdeutlicht andererseits aber auch die Tatsache, dass es *die* Lösung nicht gibt.

Der Erkenntnisgewinn der vorgelegten Arbeit besteht in Folgendem: Einerseits wird untersucht, inwiefern sich Lösungsmengen für die Faktoren X und Y angeben lassen, wann eine Faktorisierung eindeutig ist und welchen Einfluss zusätzliche Restriktionen (an die Faktoren) auf die Mengen aller möglichen Faktorisierungen haben. Andererseits ist durch den entwickelten Zerlegungsalgorithmus ein leistungsstarkes Analysewerkzeug für spektroskopische Daten bereitgestellt worden.

Aufbau der Arbeit

Die Arbeit ist wie folgt organisiert. In Kapitel 2 wird die Problemstellung detailliert erläutert. Die idealisierte Aufgabe wird als schlecht gestelltes inverses Problem charakterisiert und mittels Regularisierungen modifiziert. In Kapitel 3 wird zunächst die kontinuierliche Verallgemeinerung des diskreten multivariaten Modells (1.3) mit $u(x, y) = \sum_{i=1}^s w_i(x)v_i(y)$ und $w, v \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s)$ untersucht, wobei u bekannt und ein Faktorenpaar (w, v) gesucht ist. Das Hauptaugenmerk liegt dabei auf Untersuchungen der Faktorisierung in Bezug auf Eindeutigkeit, welche im Normalfall nicht gegeben ist. Weiter wird analysiert, unter welchen Voraussetzungen Teilfaktoren (einzelne Spalten von X oder Zeilen von Y) eindeutig sind und wie zusätzliche Restriktionen die Menge aller Faktorenpaare einschränken. Diesbezüglich stehen die Zusammenhänge zwischen den ein-

zelen Teilfaktoren $w_i(x), v_j(y)$ im Mittelpunkt. Das diskrete Faktorisierungsproblem aus (1.3) mit nichtnegativen Faktoren X, Y wird in den Kapiteln 4 und 5 untersucht. In Kapitel 6, dem numerischen Hauptteil der Arbeit, wird der neu entwickelte Algorithmus zur Berechnung einer regularisierten nichtnegativen Matrixfaktorisierung vorgestellt. Eingesetzt wird diese Lösungsmethode inzwischen unter anderem am Leibniz-Institut für Katalyse e. V. in Rostock sowie der Evonik Oxeno GmbH in Marl zur Reinkomponentenzerlegung spektroskopischer Daten. In Kapitel 7 werden letztendlich die Faktorisierungsergebnisse für eine kinetische Analyse genutzt.

Danksagung

An dieser Stelle sei den Mitarbeitern des Leibniz-Instituts für Katalyse in Rostock und der Universität Rostock Prof. A. Börner, Prof. R. Ludwig, PD Dr. D. Heller, Dr. C. Fischer, Dr. D. Selent, DC E. Barsch und DC C. Kubis sowie den Mitarbeitern der Evonik Oxeno GmbH Dr. S. Buchholz, Dr. R. Franke, Dr. D. Hess und Dr. K.-D. Wiese für eine sehr gute Zusammenarbeit und Unterstützung mein Dank ausgesprochen.

Spezieller Dank für die stets konstruktive Kritik und tatkräftige Unterstützung während der Promotionszeit ist an Prof. K. Neymeyr gerichtet.

2 Das Faktorisierungsproblem

In diesem Kapitel wird das Faktorisierungsproblem in kontinuierlicher und diskreter Gestalt vorgestellt und eine erste Untersuchung der Existenz und Eindeutigkeit einer Lösung vorgenommen. Trivialen Uneindeutigkeiten wird durch eine Einteilung in Äquivalenzklassen begegnet. Weiter wird die idealisierte Faktorisierungsaufgabe als schlecht gestelltes inverses Problem charakterisiert. Abschließend wird das multivariate Modell (1.3) aus dem (skalaren) Lambert-Beerschen Gesetz der Spektroskopie hergeleitet.

2.1 Faktordarstellungen und Faktorisierungsaufgaben

Allgemeiner Ausgangspunkt ist der folgende d -dimensionale Produktansatz. Zu $d, s \in \mathbb{N} \setminus \{0\}$, $D = D_1 \times \dots \times D_d$, $D_j \subset \mathbb{R}_+$, $j = 1, \dots, d$, wird eine Funktion $u : D \rightarrow W \subset \mathbb{R}$ der Form

$$u(x_1, \dots, x_d) = \sum_{i=1}^s \prod_{j=1}^d w_i^{(j)}(x_j) + e(x_1, \dots, x_d) \quad (2.1)$$

mit stetigen Funktionen

$$w_i^{(j)} : D_j \rightarrow \mathbb{R}_+, \quad i = 1, \dots, s, \quad j = 1, \dots, d,$$

und einer Störung $e : D_1 \times \dots \times D_d \rightarrow \mathbb{R}$ betrachtet. Weiter seien für alle $j = 1, \dots, d$ die Funktionen $w_i^{(j)}$, $i = 1, \dots, s$, jeweils auf D_j linear unabhängig.

2.1.1 Die abgeleitete Faktorisierungsaufgabe

Für die in dieser Arbeit betrachtete Aufgabenstellung ist außer der Funktion $u : D \rightarrow W$ nur bekannt, dass sich diese über einen Produktansatz wie in (2.1) zusammensetzt. Die Zahl s und die Teilfunktionen $w_i^{(j)} : D_j \rightarrow \mathbb{R}_+$, $i = 1, \dots, s$, $j = 1, \dots, d$, sind nicht bekannt. Die abgeleitete Faktorisierungsaufgabe besteht darin, zu einer geeigneten Zahl s eine Aufschlüsselung von u in die Einzelfaktoren $w_i^{(j)}$ so zu bestimmen, dass das Residuum e im Vergleich zur Faktordarstellung $\sum_{i=1}^s \prod_{j=1}^d w_i^{(j)}$ in einem geeigneten Sinne klein ist.

Die allgemeine nichtnegative Faktorisierungsaufgabe ohne Störungen

Die Betrachtung des störungsfreien Falles $e(x_1, \dots, x_d) = 0$ führt auf die idealisierte Faktorisierungsaufgabe 2.1.

Faktorisierungsaufgabe 2.1. *Zu einer stetigen Funktion $u : D \rightarrow \mathbb{R}_+$ mit $D = D_1 \times \dots \times D_d$, $D_i \subset \mathbb{R}_+$, $i = 1, \dots, d$, sind die Zahl s und eine Faktorisierung*

$$u(x_1, \dots, x_d) = \sum_{i=1}^s \prod_{j=1}^d w_i^{(j)}(x_j)$$

mit

$$w_i^{(j)} \in \mathcal{C}(D_j, \mathbb{R}_+), \quad i = 1, \dots, s, \quad j = 1, \dots, d,$$

gesucht, wobei für alle $j = 1, \dots, d$ die Funktionen $w_i^{(j)}$, $i = 1, \dots, s$, auf D_j linear unabhängig sind.

Der bivariate Ansatz

Eine Beschränkung der Dimension auf $d = 2$ führt mit der Zuordnung $y = x_2$, $x = x_1$, $v_i(y) = w_i^{(2)}(y)$, $w_i(x) = w_i^{(1)}(x)$, $i = 1, \dots, s$, für den störungsfreien Fall und $D_1 = D_2 = \mathbb{R}_+$ auf folgenden zweidimensionalen Produktansatz.

Faktorisierungsaufgabe 2.2. (Produktansatz) Zu der stetigen Funktion $u : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$ sind die Zahl $s \in \mathbb{N}$ und eine Faktorisierung

$$u(x, y) = \sum_{i=1}^s w_i(x) v_i(y) \quad (2.2)$$

in linear unabhängige Funktionen $w_i \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$, $i = 1, \dots, s$, und linear unabhängige Funktionen $v_i \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$, $i = 1, \dots, s$, gesucht.

Die diskrete Formulierung als eine nichtnegative Matrixfaktorisierung

Für eine störungsfreie Betrachtung von $u(x, y)$ auf einem Raster $\bar{x} \in \mathbb{R}^k$, $\bar{y} \in \mathbb{R}^n$,

$$A \in \mathbb{R}^{k \times n}, \quad A_{i,j} = u(\bar{x}_i, \bar{y}_j), \quad i = 1, \dots, k, \quad j = 1, \dots, n,$$

leitet sich die diskrete Formulierung einer nichtnegativen Matrixfaktorisierung ab. Dabei sind die Vektoren \bar{x}, \bar{y} so gewählt, dass A mit der Konstruktion von u wie in (2.2) den maximalen Rang $\text{rank}(A) = s$ besitzt.

Faktorisierungsaufgabe 2.3. (Nichtnegative Matrixfaktorisierung) Zu einer gegebenen Matrix $A \in \mathbb{R}_+^{k \times n}$ und $s = \text{rank}(A)$ ist eine nichtnegative Faktorisierung

$$A = XY, \quad (2.3)$$

$$X \in \mathbb{R}_+^{k \times s}, \quad Y \in \mathbb{R}_+^{s \times n}, \quad (2.4)$$

gesucht. Aufgrund von (2.3) haben X und Y den vollen Rang s .

Die Existenz einer Faktorisierung

Zu einer allgemein gewählten Funktion u ist die Existenz einer Auflösung des kontinuierlichen Produktansatzes 2.2 genauso wenig gesichert wie für eine Matrix A mit $\min\{k, n\} \geq 4$ die Existenz einer Lösung zu Faktorisierungsaufgabe 2.3. Dabei besteht im diskreten Fall die Schwierigkeit in der Beschränkung der Anzahl der Teilfaktoren (Spalten von X und Zeilen von Y) auf $s = \text{rank}(A)$, womit die Faktoren X, Y den vollen Rang s haben.

Die Frage nach der Existenz einer Auflösung der idealisierten Faktorisierungsaufgaben 2.2 und 2.3 wird in dieser Arbeit nicht untersucht. In den weiteren Analysen werden nur Probleme betrachtet, für die eine Faktorisierung exakt oder mit kleinen Störungen existiert, womit nochmals auf den Ausgangspunkt einer über (2.2) mit jeweils linear unabhängigen Funktionen $w_i(x)$, $i =$

$1, \dots, s$, und $v_i(y)$, $i = 1, \dots, s$, faktorisierbaren Funktion $u(x, y)$ hingewiesen sei. Im Speziellen heißt dies, dass einerseits Funktionen wie beispielsweise $u(x, y) = \exp(xy)$ nicht in die Analyse des störungsfreien kontinuierlichen Produktansatzes eingeschlossen werden. Ebenso werden andererseits Matrizen mit Rang s , welche sich nicht exakt in Faktoren $X \in \mathbb{R}_+^{k \times s}$, $Y \in \mathbb{R}_+^{s \times n}$ faktorisieren lassen, für Untersuchungen zur Faktorisierungsaufgabe 2.3 ausgeschlossen. Für die Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

aus Beispiel 4.22 (vergleiche auch Thomas, [84], oder Cohen, Rothblum, [13]) gilt $\text{rank}(A) = 3$, sie besitzt aber keine Faktorisierung mit $X \in \mathbb{R}_+^{4 \times 3}$, $Y \in \mathbb{R}_+^{3 \times 4}$.

Nichtlinearitätskorrekturen und Niedrigrangapproximation

Unter dem Einfluss von Störungen $e(x, y)$ aus (2.1) ändert sich für die diskrete Problemstellung das Ziel der nichtnegativen Matrixfaktorisierung zu einer nichtnegativen Niedrigrangapproximation. Weiter kann das Auftreten von Störungen negative Einträge in A hervorrufen. Erwartungsgemäß sind diese jedoch betragsmäßig klein im Vergleich zu den größten Einträgen in A .

Faktorisierungsaufgabe 2.4. (*Nichtnegative Niedrigrangapproximation*) Zu einer gegebenen Matrix $A \in \mathbb{R}^{k \times n}$, $A_{i,j} \geq -\varepsilon$ mit $0 \leq \varepsilon \ll \max_{i,j} A_{i,j}$ und einer Zahl $s \in \mathbb{N}$, $s \leq \text{rank}(A)$, ist eine nichtnegative Niedrigrangapproximation

$$\|A - XY\|_F^2 \rightarrow \min, \quad (2.5)$$

$$X \in \mathbb{R}_+^{k \times s}, \quad Y \in \mathbb{R}_+^{s \times n}, \quad (2.6)$$

vorgegebenen Ranges $\text{rank}(X) = \text{rank}(Y) = s$ gesucht. (Dabei bezeichnet $\|\cdot\|_F$ die Frobeniusnorm, siehe [22] oder Definition 4.3.)

Zusätzliche Regularisierung der Faktoren

Für das Faktorisierungsproblem 2.4 lassen sowohl die Skalierung als auch für $s \geq 2$ die Sortierung der Teilfaktoren in den Spalten von X und den Zeilen von Y keine eindeutige Auflösung zu.

In Abschnitt 4.4 wird gezeigt, dass sich die Faktorisierungsaufgabe im Normalfall auch unabhängig davon nicht eindeutig lösen lässt.

Oft basiert eine Zerlegung der Datenmatrix A auf einer Modellierung und es sind nur bestimmte Faktorisierungen zur Auswertung des Modells wertvoll. Um aus der Menge aller möglichen Faktorenpaare bestimmte zu selektieren, bedarf es einer zusätzlichen Regularisierung der einzelnen Faktoren.

Faktorisierungsaufgabe 2.5. (*Regularisierte nichtnegative Niedrigrangapproximation*) Zu einer gegebenen Matrix $A \in \mathbb{R}^{k \times n}$, $A_{i,j} \geq -\varepsilon$ mit $0 \leq \varepsilon \ll \max_{i,j} A_{i,j}$ und einer Zahl $s \in \mathbb{N}$, $s \leq \text{rank}(A)$, ist eine regularisierte nichtnegative Niedrigrangapproximation

$$\frac{1}{2} \|A - XY\|_F^2 + \frac{1}{2} \sum_{i=1}^p \gamma_i^2 f_i(X, Y) \rightarrow \min, \quad (2.7)$$

$$X \in \mathbb{R}_+^{k \times s}, \quad Y \in \mathbb{R}_+^{s \times n},$$

vorgegebenen Ranges $\text{rank}(X) = \text{rank}(Y) = s$ gesucht mit p der Anzahl der Regularisierungsfunktionen $f_i(X, Y) \geq 0$, $i = 1, \dots, p$, sowie deren Gewichtung $\gamma_i \in \mathbb{R}$, $i = 1, \dots, p$.

2.1.2 Einordnung der nichtnegativen Niedrigrangapproximation als Optimierungsproblem

Die spezielle Einordnung der Faktorisierungsaufgaben 2.3 und 2.4 ist die nichtnegative Matrixfaktorisierung, siehe beispielsweise [3, 12, 47, 58, 75].

Für die allgemeine Einordnung von 2.4 als Optimierungsproblem ist die fehlende Konvexität der Zielfunktion aus (2.5) entscheidend, was für die Lösung der zugehörigen Faktorisierungsaufgabe nicht von Vorteil ist. Die Funktion

$$f(X, Y) = \|A - XY\|_F^2 \quad (2.8)$$

enthält bei einer simultanen Betrachtung der Unbekannten X und Y Terme vierten Grades in $X_{i,j}$ und $Y_{l,m}$ und ist nicht konvex. Das Optimierungsproblem fällt unter den Nebenbedingungen $X, Y \geq 0$ somit in die Klasse der nichtkonvexen Optimierungsprobleme unter Nichtnegativitätsnebenbedingungen, vergleiche auch Bemerkung 5.2.

Einen Ausweg bietet der beispielsweise von Kim und Park in [47] betrachtete Zugang, eine Faktorisierung über die alternierende Lösung linearer Ausgleichsprobleme unter Nichtnegativitätsrestriktionen zu berechnen. Unter Fixierung eines Faktors (X beziehungsweise Y) werden alternierend Optimierungsprobleme aufgestellt und in Y beziehungsweise X gelöst. Das hat jedoch den Verlust der simultanen Auflösung der Faktoren zur Folge. Diese und weitere klassische Vorgehensweisen zur Berechnung von Faktorisierungen sind in Kapitel 5 detailliert vorgestellt.

2.2 Die Faktorisierungsaufgabe als schlecht gestelltes Problem

2.2.1 Inverse Probleme und deren Klassifizierung

Die gestellten Faktorisierungsaufgaben sind in die Klasse der inversen Probleme einzuordnen. Seien \mathcal{X} und \mathcal{Y} zwei beliebige Mengen. Zu der Operatorgleichung

$$F(x) = y, \quad x \in D \subseteq \mathcal{X}, \quad y \in \mathcal{Y},$$

liegt ein *inverses Problem* vor, falls zu einem Element $y \in \mathcal{Y}$ ein $x \in D$ gesucht ist mit $F(x) = y$.

Die folgende von Hadamard eingeführte Charakterisierung inverser Probleme unterscheidet korrekt und schlecht gestellte Aufgaben, [36].

Definition 2.6. *Die Operatorgleichung*

$$F(x) = y, \quad x \in D \subseteq \mathcal{X}, \quad y \in \mathcal{Y},$$

und damit das inverse Problem heißt nach Hadamard korrekt gestellt, falls gilt:

1. Zu jedem $y \in \mathcal{Y}$ gibt es eine Lösung $x \in D$ von $F(x) = y$ (Existenzbedingung).
2. Die aus $F(x) = y$ erhaltene Lösung x ist in D eindeutig bestimmt (Eindeutigkeit).
3. Die Lösung x hängt stetig von der rechten Seite y ab (Stabilitätsbedingung).

Andernfalls heißt die Aufgabe schlecht gestellt.

2.2.2 Das inverse Problem des kontinuierlichen Produktansatzes

Bei der Klassifizierung des kontinuierlichen Produktsatzes aus Faktorisierungsaufgabe 2.2 nach Definition 2.6 ergeben sich unmittelbar zwei Probleme. Einerseits besitzt nicht jede nicht-negative Funktion eine Auflösung wie in (2.2) und andererseits stellt sich für den Fall der Existenz einer Faktorisierung die Frage nach deren Eindeutigkeit.

Für die in dieser Arbeit betrachtete Aufgabenstellung ist speziell die Frage der Eindeutigkeit von Interesse. Um die Skalierung und Sortierung der Faktoren unbeachtet zu lassen, ist es diesbezüglich sinnvoll eine Einteilung der Faktorisierungen in Äquivalenzklassen vorzunehmen.

Nichteindeutigkeit der Faktorisierung durch Skalierung und Permutierung: Einführung von Äquivalenzklassen

Seien (w, v) und (\tilde{w}, \tilde{v}) mit $w, v, \tilde{w}, \tilde{v} \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s)$ zwei nichtnegative Faktorisierungen zum kontinuierlichen Produktansatz

$$u(x, y) = \sum_{i=1}^s w_i(x) v_i(y) = \sum_{i=1}^s \tilde{w}_i(x) \tilde{v}_i(y).$$

In Lemma 3.5 wird später gezeigt, dass es eine reguläre Transformation $T \in \mathbb{R}^{s \times s}$ gibt mit

$$\tilde{w}_i(x) = \sum_{j=1}^s w_j(x) (T^{-1})_{j,i}, \quad \tilde{v}_i(y) = \sum_{j=1}^s T_{i,j} v_j(y), \quad i = 1, \dots, s. \quad (2.9)$$

Für den Umgang mit der freien Skalierung und der freien Sortierung einer Faktorisierung wird folgende Relation in Definition 2.7 eingeführt. Es ist dazu mit $s \in \mathbb{N}$ die Menge

$$\mathcal{X}^{(s)} = \{(w, v) : w, v \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s), w_1(x), \dots, w_s(x) \text{ lin. unabh.}, v_1(y), \dots, v_s(y) \text{ lin. unabh.}\}$$

definiert.

Definition 2.7. Zu einer gegebenen und wie in (2.2) faktorisierbaren Funktion $u \in \mathcal{C}(\mathbb{R}_+^2, \mathbb{R}_+)$ sei auf der Menge $\mathcal{X}^{(s)}$ die Relation R wie folgt definiert

$$R \subset \mathcal{X}^{(s)} \times \mathcal{X}^{(s)},$$

$$(w, v) \sim_R (\tilde{w}, \tilde{v}) : \Leftrightarrow u(x, y) = \sum w_i(x) v_i(y) = \sum \tilde{w}_i(x) \tilde{v}_i(y) \quad \text{und}$$

$$T \in \mathbb{R}^{s \times s} \text{ aus (2.9) ist eine permutierte Diagonalmatrix.}$$

Korollar 2.8. Die Relation R ist reflexiv, symmetrisch und transitiv und somit eine Äquivalenzrelation. Es lassen sich die Faktorisierungen von u durch R in Äquivalenzklassen

$$[(w, v)]_R$$

unterteilen.

2.2.3 Die Einordnung als schlecht gestelltes Problem

Die Frage nach der Eindeutigkeit der Faktorisierung ist ein Teil der Untersuchungen in Kapitel 3. Dabei stellt sich heraus, dass das Faktorisierungsproblem 2.2 für $s > 1$ in der Regel mehr als eine Äquivalenzklasse von Auflösungen besitzt. Eigenschaft 2 von Definition 2.6 ist selbst

bei einer Auflösung in Äquivalenzklassen nicht erfüllt und das inverse Problem aus Faktorisierungsaufgabe 2.2 ist schlecht gestellt. Gleiches gilt für die nichtnegative Matrixfaktorisierung aus Faktorisierungsaufgabe 2.3, vergleiche Abschnitt 4.4.

Es besteht der Bedarf, aus der Menge aller möglichen Faktorisierungen bestimmte Äquivalenzklassen von Faktorenpaaren zu selektieren. Damit ist die Faktorisierungsaufgabe 2.5 gestellt; zu A ist eine regularisierte nichtnegative Niedrigrangapproximation gesucht.

2.3 Spektroskopie, Chemometrie und approximative Faktorisierungen

„CHEMOMETRICS, IN THE MOST GENERAL SENSE, IS THE ART OF PROCESSING DATA WITH VARIOUS NUMERICAL TECHNIQUES IN ORDER TO EXTRACT USEFUL INFORMATION“, R. Kramer, [50].

2.3.1 Das Gesetz von Lambert-Beer

In der Spektroskopie ist die Absorption α beim Strahlendurchgang durch lichtabsorbierende Materie definiert als der dekadische Logarithmus des Quotienten der Intensität der einfallenden Strahlung I_0 und der ausfallenden I_1 . Die molare Konzentration \hat{c} der Probe, deren Extinktionskoeffizient \hat{a} und α stehen in folgender Beziehung.

Das Gesetz von Lambert und Beer, [28, 29],

$$\alpha = \alpha(\lambda) = \log_{10} \frac{I_0}{I_1} = \hat{c} \hat{a}(\lambda) l \quad (2.10)$$

beschreibt den linearen Zusammenhang zwischen der Absorption α einer chemischen Lösung, deren molaren dekadischen Extinktionskoeffizienten (Absorptivität) \hat{a} bei der Wellenlänge λ , deren molarer Konzentration \hat{c} und der Probendicke l .

2.3.2 Superposition der Komponenten

Bei der Untersuchung eines Systems aus s Komponenten ergibt sich die Gesamtabsorption in Gleichung (2.10) als Superposition der Einzelabsorptionen

$$\alpha(\lambda) = l \sum_{i=1}^s \hat{c}_i \hat{a}_i(\lambda).$$

Wird die Probenabsorption zur Analyse chemischer Reaktionsverläufe genutzt, bestimmt sich die Probenabsorption α abhängig von zwei veränderlichen Größen. Neben den Extinktionskoeffizienten $\hat{a}_i = \hat{a}_i(\lambda)$ in Abhängigkeit von der Wellenlänge sind auch die Konzentrationen $\hat{c}_i = \hat{c}_i(t)$ als Funktionen in Abhängigkeit von der Zeit maßgebend. Ohne wesentliche Einschränkung darf die Schichtdicke l auf 1 normiert werden. Idealisiert lässt sich die Absorption als multivariate reellwertige Funktion $u : \mathbb{R}_+^2 \rightarrow \mathbb{R}_+$,

$$u(t, \lambda) = \sum_{j=1}^s \hat{c}_j(t) \hat{a}_j(\lambda) \quad (2.11)$$

angeben.

2.3.3 Funktionsauswertungen auf einem Zeit- und Frequenzraster und Fehlerterme

In praktischen Untersuchungen stehen für die Funktion u aus (2.11) nur endlich viele Messwerte auf einem Raster $\bar{t} \in \mathbb{R}^k$ und $\bar{\lambda} \in \mathbb{R}^n$ zur Verfügung. Dies führt auf die Matrizen $A \in \mathbb{R}_+^{k \times n}$, $\hat{C} \in \mathbb{R}_+^{k \times s}$, $\hat{A} \in \mathbb{R}_+^{s \times n}$,

$$A_{i,j} = u(\bar{t}_i, \bar{\lambda}_j), \quad \hat{C}_{i,l} = \hat{c}_l(\bar{t}_i), \quad \hat{A}_{l,j} = \hat{a}_l(\bar{\lambda}_j), \quad i = 1, \dots, k, \quad j = 1, \dots, n, \quad l = 1, \dots, s,$$

und das Produkt $A = \hat{C}\hat{A}$. Dabei ist \hat{C} die Matrix der Konzentrationsprofile und \hat{A} die Absorptivitätsmatrix. Die Modellannahme beinhaltet Störungen, die ausgeglichen werden müssen

$$A = \hat{C}\hat{A} + E, \quad E \in \mathbb{R}^{k \times n}. \quad (2.12)$$

Die Einträge in E basieren einerseits auf systematischen Modellfehlern in Form von Nichtlinearitätskorrekturen des idealisierten Gesetzes von Lambert-Beer sowie andererseits auf zufälligen Fehlern. Eine detaillierte Erläuterung zur Zusammensetzung der Störungen bei spektroskopischen Daten erfolgt im Anhang.

2.3.4 Die Faktorisierungsaufgabe

Mit der Grundlage des Lambert-Beerschen Gesetzes in der Matrixformulierung (2.12) und unter der Annahme $E = 0$ stellt sich die Gleichung $A = \hat{C}\hat{A}$ als Matrixfaktorisierungsaufgabe dar. Ausgehend von der Matrix der Messdaten sollen die Faktoren \hat{C} und \hat{A} so bestimmt werden, dass deren Produkt möglichst gut A approximiert. Zugleich sind die Matrixelemente von \hat{C} und \hat{A} aus physikalischen Gründen (Konzentration, Absorptivität) nichtnegativ. Damit ist also eine nichtnegative Niedrigrangapproximation von A gesucht. Die Lösung dieses Problems ist im Normalfall nicht eindeutig; die Aufgabenstellung wird zu einer regularisierten nichtnegativen Niedrigrangapproximation modifiziert.

3 Die kontinuierliche Faktorisierungsaufgabe

In diesem Kapitel erfolgt eine Analyse des kontinuierlichen Produktansatzes. Dabei wird der Problemstellung aus den Abschnitten 2.1 und 2.3 entsprechend eine Funktion $u(x, y)$ zugrunde gelegt, welche sich über (2.1) mit $d = 2$ und $e(x, y) = 0$ aus s Faktorenpaaren zusammensetzt; es werden demzufolge nur Funktionen untersucht, zu denen eine Faktorisierung existiert. Es bleibt die Frage nach der Eindeutigkeit der Auflösung.

Im Vordergrund steht die Charakterisierung der Faktorisierung bezüglich der Anzahl der Äquivalenzklassen von Auflösungen in „eindeutig lösbar“ und „mehrdeutig lösbar“. Dazu werden Zusammenhänge unter möglichen Faktorisierungen untersucht und im Zuge dessen ein notwendiges und ein hinreichendes Kriterium dafür formuliert, dass alle nichtnegativen Zerlegungen äquivalent sind. Dabei werden bereits die notwendigen Voraussetzungen nur in seltenen Fällen erfüllt.

Weiter wird analysiert, unter welchen Umständen einzelne Teilfaktoren eindeutig sind. Insbesondere steht im Fokus, inwiefern derartige Teileindeutigkeiten (oder analog dazu zusätzliche Restriktionen) die Vielfalt aller Faktorisierungen einschränken; es wird untersucht, wie sich die Eindeutigkeit von Teilfaktoren $w_i(x)$ oder $v_j(y)$ indirekt auf andere Teilfaktoren auswirkt. Dafür sind speziell die Zusammenhänge zwischen den einzelnen Teilfaktoren $w_i(x), v_j(y)$, $i, j = 1, \dots, s$, und die Nichtnegativität aller Teilfaktoren entscheidend.

3.1 Zusammenhänge zwischen den Faktorisierungen und ihre Einteilung in Äquivalenzklassen

Die Grundlage von Faktorisierungsaufgabe 2.2 bildet mit $s \in \mathbb{N}$, $w, v \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s)$ und $u \in \mathcal{C}(\mathbb{R}_+^2, \mathbb{R}_+)$ ein Zusammenhang

$$u(x, y) = \sum_{i=1}^s w_i(x) v_i(y). \quad (3.1)$$

Dabei ist jeweils die lineare Unabhängigkeit der s Funktionen $w_i(x)$, $i = 1, \dots, s$, beziehungsweise $v_i(y)$, $i = 1, \dots, s$, zu beachten, das heißt es existieren zwei s -Tupel $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ mit

$$\det(U) \neq 0, \quad U \in \mathbb{R}^{s \times s}, \quad U_{i,j} = u(\bar{x}_i, \bar{y}_j), \quad i, j = 1, \dots, s,$$

und damit auch

$$\det(X), \det(Y) \neq 0, \quad X, Y \in \mathbb{R}^{s \times s}, \quad X_{i,j} = w_j(\bar{x}_i), \quad Y_{i,j} = v_i(\bar{y}_j), \quad i, j = 1, \dots, s.$$

Als Vereinfachung für diesen Abschnitt werden folgende Voraussetzung und folgende Bezeichnung definiert. Weiter wird nur der Fall $s \geq 2$ betrachtet.

Faktorisierungsvoraussetzung 3.1. *Gegeben seien eine Funktion $u \in \mathcal{C}(\mathbb{R}_+^2, \mathbb{R}_+)$, welche sich wie in (3.1) mit $s \geq 2$ als Produkt aus linear unabhängigen Funktionen $w_i \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$, $i = 1, \dots, s$, und linear unabhängigen Funktionen $v_i \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+)$, $i = 1, \dots, s$, zusammensetzt.*

Bezeichnung 3.2. Zu der Faktorisierungsvoraussetzung 3.1 sowie zwei Vektoren $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ mit $\text{rank}(U) = s$, $U \in \mathbb{R}^{s \times s}$, $U_{i,j} = u(\bar{x}_i, \bar{y}_j)$ sind die Funktionen $u_{\bar{y}}, u_{\bar{x}} \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s)$ definiert als

$$\begin{aligned} u_{\bar{y}}(x) &= (u(x, \bar{y}_1), \dots, u(x, \bar{y}_s))^T, \\ u_{\bar{x}}(y) &= (u(\bar{x}_1, y), \dots, u(\bar{x}_s, y))^T. \end{aligned}$$

3.1.1 Das inverse Problem

Wie in Faktorisierungsaufgabe 2.2 formuliert, ist in der Umkehrung aus (3.1) nur die Funktion $u(x, y)$ bekannt. Die inverse Aufgabe ist es, zu u die Zahl $s \in \mathbb{N}$ sowie eine nichtnegative Faktorisierung in jeweils untereinander linear unabhängige Funktionen $w_i : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, $i = 1, \dots, s$, und $v_i : \mathbb{R}_+ \rightarrow \mathbb{R}_+$, $i = 1, \dots, s$, zu bestimmen.

Die Zahl s lässt sich mittels u und unter Verwendung der linearen Unabhängigkeit der s Teilfunktionen aus $w(x)$ und $v(y)$ untereinander bestimmen.

Lemma 3.3. Es sei die Faktorisierungsvoraussetzung 3.1 gegeben.

Die Zahl $s \in \mathbb{N}$ ist die größte Zahl, für die es zwei Tupel $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ gibt mit

$$\det(U) \neq 0, \quad U \in \mathbb{R}^{s \times s}, \quad U_{i,j} = u(\bar{x}_i, \bar{y}_j), \quad i, j = 1, \dots, s,$$

also

$$s = \lim_{m,n \rightarrow \infty} \max_{\substack{\bar{x} \in \mathbb{R}_+^m, \\ \bar{y} \in \mathbb{R}_+^n}} \text{rank}((u(\bar{x}_i, \bar{y}_j))_{i=1,\dots,m, j=1,\dots,n}). \quad (3.2)$$

Durch die Konstruktion von u über (3.1) ist s klar definiert.

Mögliche Basen der Funktionen $w_i(x)$, $i = 1, \dots, s$, und $v_i(y)$, $i = 1, \dots, s$, sind bei geeigneter Wahl von $\bar{x}, \bar{y} \in \mathbb{R}_+^s$, das heißt $\text{rank}(U) = s$,

$$\begin{aligned} B_X &= \{u(x, \bar{y}_1), \dots, u(x, \bar{y}_s)\}, \\ B_Y &= \{u(\bar{x}_1, y), \dots, u(\bar{x}_s, y)\}. \end{aligned}$$

Beweis: Offensichtlich gilt (3.2). Bleibt zunächst zu zeigen, dass B_X eine Basis für $w(x)$ ist. Seien $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ so gewählt, dass $\text{rank}(U) = s$ gilt. Weiter seien durch $\tilde{w}, \tilde{v} \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s)$ eine beliebige nichtnegative Faktorisierung von u gegeben und $\tilde{Y} \in \mathbb{R}^{s \times s}$ definiert als $\tilde{Y}_{i,j} = \tilde{v}_i(\bar{y}_j)$, $i, j = 1, \dots, s$. Aus $\text{rank}(U) = s$ folgt $\text{rank}(\tilde{Y}) = s$. Dann ist mit der Bezeichnung 3.2 der Faktor $\tilde{w}(x)$ Lösung des linearen Gleichungssystems $(u_{\bar{y}}(x))^T = (\tilde{w}_1(x), \dots, \tilde{w}_s(x))\tilde{Y}$, also

$$(\tilde{w}(x))^T = (u_{\bar{y}}(x))^T \tilde{Y}^{-1},$$

und B_X ist ein Erzeugendensystem der Faktoren $w_i(x)$, $i = 1, \dots, s$. Da sich $u(x, y)$ aus s Teilfaktoren zusammensetzt, ist B_X eine Basis. Analog lässt sich zeigen, dass B_Y eine Basis für $v_i(y)$, $i = 1, \dots, s$, ist. \square

Korollar 3.4. Seien wie in Lemma 3.3 die Tupel $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ geeignet gewählt mit $\text{rank}(U) = s$, $U \in \mathbb{R}^{s \times s}$, $U_{i,j} = u(\bar{x}_i, \bar{y}_j)$. Dann lässt sich die Aufgabenstellung auf die Bestimmung geeigneter Matrizen $S, T \in \mathbb{R}^{s \times s}$

$$\begin{aligned} w_i(x) &= \sum_{j=1}^s u(x, \bar{y}_j) S_{j,i}, & i = 1, \dots, s, \\ v_i(y) &= \sum_{j=1}^s T_{i,j} u(\bar{x}_j, y), & i = 1, \dots, s, \end{aligned}$$

reduzieren.

3.1.2 Zusammenhänge unter den Faktorisierungen

Es stellt sich unmittelbar die Frage nach der Eindeutigkeit der Zerlegung. Dazu wird zunächst in einer Vorüberlegung der Zusammenhang zwischen zwei Faktorisierungen untersucht. Anschließend werden alle nichtnegativen Faktorisierungen über die Relation aus Definition 2.7 in Äquivalenzklassen eingeteilt.

Lemma 3.5. *Es sei die Faktorisierungsvoraussetzung 3.1 gegeben.*

Dann hat jede weitere Zerlegung von $u(x, y)$

$$u(x, y) = \sum_{i=1}^s \tilde{w}_i(x) \tilde{v}_i(y)$$

die Gestalt

$$\tilde{w}_i(x) = \sum_{j=1}^s w_j(x) (T^{-1})_{j,i}, \quad i = 1, \dots, s, \quad (3.3a)$$

$$\tilde{v}_i(y) = \sum_{j=1}^s T_{i,j} v_j(y), \quad i = 1, \dots, s, \quad (3.3b)$$

mit einer regulären Matrix $T \in \mathbb{R}^{s \times s}$.

Beweis: *Durch Lemma 3.3 ist klar, dass es Matrizen $S, T \in \mathbb{R}^{s \times s}$ mit*

$$\tilde{w}_i(x) = \sum_{j=1}^s w_j(x) S_{j,i} \quad \tilde{v}_i(y) = \sum_{j=1}^s T_{i,j} v_j(y)$$

gibt. Dass diese regulär sind, folgt aus $u(x, y) = \sum w_i(x) v_i(y) = \sum \tilde{w}_i(x) \tilde{v}_i(y)$ und der Faktorisierungsvoraussetzung 3.1. Es bleibt $S = T^{-1}$ zu zeigen.

Seien $\bar{x}, \bar{y} \in \mathbb{R}^s$ derart gewählt, dass für $U \in \mathbb{R}^{s \times s}$, $U_{i,j} = u(\bar{x}_i, \bar{y}_j)$, gilt $\text{rank}(U) = s$. Weiter seien $X, \tilde{X}, Y, \tilde{Y} \in \mathbb{R}^{s \times s}$, $X_{i,j} = w_j(\bar{x}_i)$, $\tilde{X}_{i,j} = \tilde{w}_j(\bar{x}_i)$, $Y_{i,j} = v_i(\bar{y}_j)$, $\tilde{Y}_{i,j} = \tilde{v}_i(\bar{y}_j)$, $i, j = 1, \dots, s$. Aus $\text{rank}(U) = s$ folgt $\text{rank}(X) = \text{rank}(Y) = \text{rank}(\tilde{X}) = \text{rank}(\tilde{Y}) = s$. Aus der Faktorisierungsvoraussetzung ergibt sich

$$\tilde{X} \tilde{Y} = X S T Y$$

und es folgt, dass $S = T^{-1}$ gelten muss. □

3.1.3 Die Einteilung in Äquivalenzklassen

Um in Bezug auf die Eindeutigkeit der Faktorisierung Skalierung und Permutierung für die einzelnen Faktoren unbeachtet zu lassen, werden alle Faktorisierungen über die Relation R aus Definition 2.7 in Äquivalenzklassen eingeteilt.

Nach Definition 2.7 gilt für zwei nichtnegative Faktorisierungen (w, v) und (\tilde{w}, \tilde{v}) von u

$$(w(x), v(y)) \sim_R (\tilde{w}(x), \tilde{v}(y)) : \Leftrightarrow \exists T \in \mathbb{R}^{s \times s} \text{ regulär mit genau } s \text{ Nichtnullelementen,} \\ \text{so dass (3.3a) und (3.3b) erfüllt sind.}$$

3.2 Nichteindeutigkeit der Faktorisierung

„THE FOLLOWING RULE OF THUMB TELLS WHEN TO EXPECT A UNIQUE FACTORIZATION: IF ALL COLUMNS OF THE CORRECT UNKNOWN G AND ALL ROWS OF THE CORRECT UNKNOWN F CONTAIN A SIGNIFICANT NUMBER OF ZEROS, THEN THE RESULT BY PMF IS UNIQUE.“, Paatero, Tapper, [67], für die diskrete Aufgabenstellung einer nichtnegativen Matrixfaktorisierung $A = GF$. (PMF: „positive matrix factorization“)

Im vorherigen Abschnitt wurde in Lemma 3.5 die Transformationseigenschaft der Faktorisierungen untersucht. Dies führt direkt auf die Eindeutigkeitsproblematik und die Frage, unter welchen Voraussetzungen alle möglichen Faktorisierungen in derselben Äquivalenzklasse liegen. Dazu werden in diesem Abschnitt in Satz 3.6 notwendige und in Satz 3.8 hinreichende Kriterien aufgestellt.

3.2.1 Notwendige Voraussetzung für eine eindeutige Faktorisierung

Als notwendige Bedingung für eine eindeutige Faktorisierung ergibt sich aus der Faktorisierungsforderung und der Nichtnegativität der Faktoren folgendes Kriterium. Entscheidender Punkt ist dabei, dass es keine zwei Funktionen w_i, w_j mit $w_i(x) \geq \alpha w_j(x)$ für ein $\alpha > 0$ sowie alle $x \in \mathbb{R}_+$ gibt. Dies würde die Subtraktion $\tilde{w}_i = w_i - \alpha w_j$ ermöglichen, was auf eine andere nichtnegative Faktorisierung führt.

Satz 3.6. (Notwendige Bedingung) Seien die Faktorisierungsvoraussetzung 3.1 erfüllt und die nichtnegative Faktorisierung von $u(x, y)$ durch Funktionen w_i, v_i , $i = 1, \dots, s$, bis auf Skalierung und Sortierung eindeutig.

Dann gibt es für alle $i, j \in \{1, \dots, s\}$, $i \neq j$, Stellen $x, y \in \mathbb{R}_+$ mit

$$\begin{aligned} w_i(x) &= 0 < w_j(x), \\ v_i(y) &= 0 < v_j(y), \end{aligned}$$

was im Umkehrschluß heißt, dass es keine $i, j \in \{1, \dots, s\}$ und kein $\alpha > 0$ gibt mit $i \neq j$ und

$$w_i(x) \leq \alpha w_j(x), \quad \forall x \geq 0, \quad (3.4a)$$

oder

$$v_i(y) \leq \alpha v_j(y), \quad \forall y \geq 0. \quad (3.4b)$$

Für den Fall $s = 2$ ist diese Bedingung gleichzeitig hinreichend, siehe Satz 3.8.

Beweis: Es wird nur die Forderung aus (3.4b) gezeigt, die Bedingung (3.4a) für $w(x)$ folgt analog. Sei angenommen die Faktorisierung von u ist eindeutig und es gebe ein $\alpha > 0$ sowie i, j mit $i \neq j$ und $v_i(y) \leq \alpha v_j(y)$ für alle $y \geq 0$. Die Konstruktion von $\tilde{w}(x)$ und $\tilde{v}(y)$ als

$$\begin{aligned} \tilde{v}_l &= v_l, \quad l \in \{1, \dots, s\} \setminus \{j\}, & \tilde{w}_l &= w_l, \quad l \in \{1, \dots, s\} \setminus \{i\}, \\ \tilde{v}_j &= v_j - \alpha^{-1} v_i, & \tilde{w}_i &= w_i + \alpha w_j, \end{aligned}$$

ist wegen $s \geq 2$ möglich. Die Funktionen \tilde{w} und \tilde{v} erfüllen die Nichtnegativitätsbedingungen, rekonstruieren $u(x, y)$ und liegen bezüglich der Relation R aus 2.7 nicht in derselben Äquivalenzklasse wie $(w(x), v(y))$. Die Transformation T aus (3.3a) und (3.3b) ist dabei

$$T = I_{s \times s} - \alpha^{-1} e_j e_i^T.$$

Damit ist die Faktorisierung nicht eindeutig, was der Voraussetzung widerspricht. \square

3.2.2 Hinreichende Voraussetzung für eine eindeutige Faktorisierung

Bezüglich hinreichender Voraussetzungen für eine eindeutige Faktorisierung liefert Satz 3.8 eine Aussage. Dazu wird Lemma 3.7 benötigt. Dieses zeigt, dass eine reguläre, nichtnegative Matrix $T \in \mathbb{R}_+^{s \times s}$ entweder genau s Nichtnullelemente enthält oder eine Inverse mit einem negativen Eintrag besitzt. Dies ist für mögliche Transformationen einer Faktorisierung entscheidend um nichtnegative Funktionen zu erhalten.

Lemma 3.7. *Gegeben sei eine reguläre Matrix $T \in \mathbb{R}^{s \times s}$ mit $T_{i,j} \geq 0$, $i, j = 1, \dots, s$. Weiter seien $1 \leq s_0, i_0, i_1 \leq s$ und $i_0 \neq i_1$ mit*

$$T_{s_0, i_0}, T_{s_0, i_1} > 0.$$

Dann gibt es s_1 und j_0 mit

$$(T^{-1})_{s_1, j_0} < 0.$$

Beweis: *Angenommen es gelte $(T^{-1})_{i,j} \geq 0$ für alle $i, j = 1, \dots, s$. Es gilt aber*

$$(T T^{-1})_{s_0, j} = \sum_{i=1}^s T_{s_0, i} (T^{-1})_{i, j} = 0, \quad \forall j \neq s_0,$$

woraus mit $(T^{-1})_{i,j} \geq 0$, $i, j = 1, \dots, s$, folgt, dass

$$\begin{aligned} (T^{-1})_{i_0, j} &= (T^{-1})_{i_1, j} = 0, \\ \Rightarrow (T^{-1})_{i_0, :} &= \alpha (T^{-1})_{i_1, :} \Rightarrow \det(T^{-1}) = 0, \end{aligned} \quad \forall j \neq s_0,$$

was der Regularität von T widerspricht. □

Satz 3.8. *(Hinreichende Bedingung) Es sei die Faktorisierungsvoraussetzung 3.1 gegeben. Weiter sei (w, v) eine nichtnegative Faktorisierung von u und es gebe $\bar{x}, \bar{y}, \epsilon, \varepsilon \in \mathbb{R}_+^s$, $\epsilon_i, \varepsilon_i > 0$, $i = 1, \dots, s$, mit*

$$w_i(\bar{x}_j) = \epsilon_i \delta_{i,j}, \quad \forall i, j = 1, \dots, s, \quad (3.5a)$$

$$v_i(\bar{y}_j) = \varepsilon_i \delta_{i,j}, \quad \forall i, j = 1, \dots, s, \quad (3.5b)$$

und dem Kronecker-Delta $\delta_{i,j}$.

Dann sind alle nichtnegativen Faktorisierungen von u äquivalent.

Beweis: *Mittels (3.5a) ergibt sich für jede weitere nichtnegative Faktorisierung $(\tilde{w}(x), \tilde{v}(y))$ eine Transformationsmatrix T^{-1} aus (3.3a) als nichtnegativ. Ebenso folgt aus (3.5b) und (3.3b), dass T nichtnegativ ist. Die Matrix T ist regulär und kann damit nach Lemma 3.7 nur s Elemente größer Null besitzen, damit liegen $(w(x), v(y))$ und $(\tilde{w}(x), \tilde{v}(y))$ in einer Äquivalenzklasse.* □

Weiter stellt sich durch die zusammengetragenen Bedingungen bezüglich der Eindeutigkeit die Frage, wie anhand von $u(x, y)$ entschieden werden kann, ob eine Zerlegung eindeutig ist.

Bemerkung 3.9. *Einerseits kann, sofern eine Zerlegung bekannt ist, unter Anwendung von Satz 3.8 mittels dieser entschieden werden, ob das Problem nur eine Äquivalenzklasse an Auflösungen besitzt oder nicht. Andererseits kann eine Funktion als eindeutig faktorisierbar deklariert werden, falls es $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ und $\varepsilon \in (\mathbb{R}_+ \setminus \{0\})^s$ gibt mit*

$$u(\bar{x}_i, \bar{y}_j) = \varepsilon_i \delta_{i,j}.$$

Als Faktorisierung ergibt sich dabei unmittelbar

$$w(x) = u_{\bar{y}}(x), \quad v(y) = \text{diag}(\varepsilon_1^{-1}, \dots, \varepsilon_s^{-1}) u_{\bar{x}}(y).$$

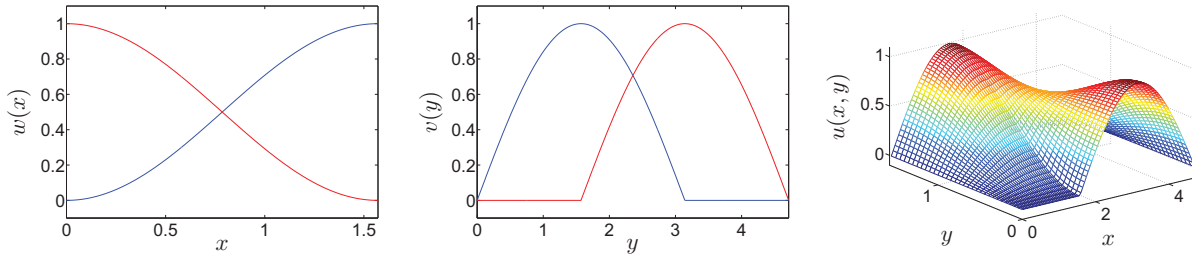


Abbildung 3.1: Dargestellt sind die Funktionen $w(x)$, $v(y)$ und $u(x, y)$ aus Beispiel 3.10. Die hinreichenden Voraussetzungen für eine eindeutige Faktorisierung sind erfüllt; es lassen sich keine nichtnegativen Differenzen unter den einzelnen Teilfaktoren bilden. Alle Faktorisierungen liegen in derselben Äquivalenzklasse.

Zur Veranschaulichung der notwendigen und hinreichenden Kriterien sollen die beiden folgenden Beispiele dienen.

Beispiel 3.10. (Eindeutige Faktorisierung) Es sei eine Faktorisierung aus $s = 2$ Teilfaktoren gegeben. Dazu seien die Intervalle $x \in [0, \frac{1}{2}\pi]$ und $y \in [0, \frac{3}{2}\pi]$ und die Faktoren

$$w(x) = \begin{pmatrix} \sin^2(x) \\ \cos^2(x) \end{pmatrix}, \quad v(y) = \begin{pmatrix} \max(\sin(y), 0) \\ \max(\sin(y - \frac{1}{2}\pi), 0) \end{pmatrix}$$

gewählt. Die Faktoren und die resultierende Funktion $u(x, y) = (w(x))^T v(y)$ sind in Abbildung 3.1 dargestellt.

Die Faktoren erfüllen die Voraussetzungen aus Satz 3.8. Somit liegen alle Faktorisierungen von $u(x, y)$ in derselben Äquivalenzklasse.

Beispiel 3.11. (Keine eindeutige Faktorisierung) Etwas abgewandelt zu Beispiel 3.10 ist folgendes Problem. Als Intervalle sind erneut $x \in [0, \frac{1}{2}\pi]$ und $y \in [0, \frac{3}{2}\pi]$ gewählt. Die Faktoren

$$w(x) = \begin{pmatrix} \sin^2(x) \\ \frac{1}{4}\sin^2(2x) \end{pmatrix}, \quad v(y) = \begin{pmatrix} \max(\sin(y), 0) \\ \max(\sin(y - \frac{1}{2}\pi), 0) \end{pmatrix}$$

erfüllen die notwendige Voraussetzung aus Satz 3.6 nicht. Die Differenz $w_1(x) - \alpha w_2(x)$ ist für $\alpha \in [0, 1]$ auf dem Intervall $x \in [0, \frac{1}{2}\pi]$ nichtnegativ. Es lassen sich mit

$$(\hat{w}_\alpha(x))^T = (w(x))^T \begin{pmatrix} 1 & 0 \\ -\alpha & 1 \end{pmatrix}, \quad \hat{v}_\alpha(y) = \begin{pmatrix} 1 & 0 \\ \alpha & 1 \end{pmatrix} v(y)$$

weitere nichtnegative Faktorisierungen von $u = w^T v$ bilden, siehe dazu auch Abbildung 3.2.

Anzumerken ist, dass alle nichtnegativen Faktorisierungen unskaliert die Teilfaktoren $w_2(x)$ und $v_1(y)$ enthalten. Die Zusammenhänge dazu werden später in den Sätzen 3.15 und 3.19 untersucht.

3.2.3 Abschwächung der hinreichenden Bedingung

Die hinreichenden Voraussetzungen an $u(x, y)$ für nur eine Äquivalenzklasse an Faktorisierungen lassen sich geringfügig abschwächen. Dazu sei zunächst das Lemma 3.12 gezeigt sowie anschließend Satz 3.13 formuliert.

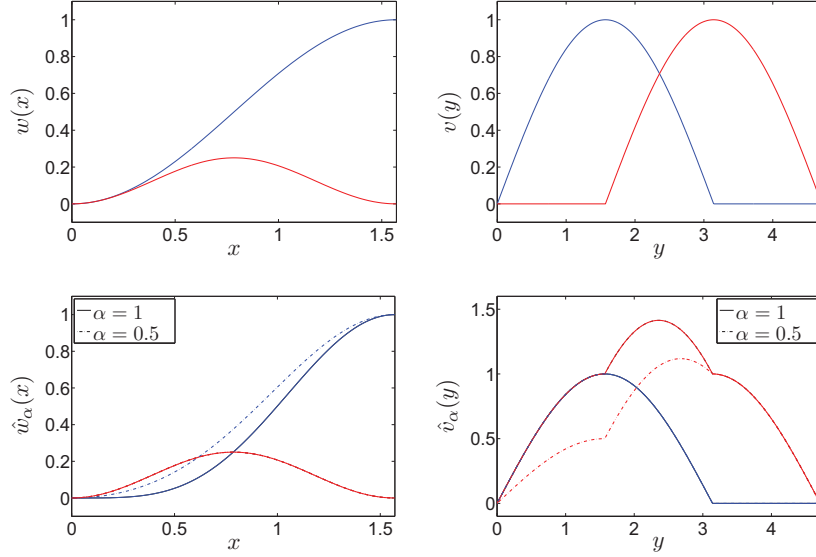


Abbildung 3.2: Dargestellt sind oben die Funktionen w , v aus Beispiel 3.11. Die notwendigen Voraussetzungen für eine eindeutige Faktorisierung sind nicht erfüllt; es existieren mehrere Äquivalenzklassen von Faktorisierungen. Zwei zu (w, v) alternative Lösungen sind unten mit $(\hat{w}_\alpha, \hat{v}_\alpha)$ für $\alpha = 0.5$ und $\alpha = 1$ dargestellt. Anmerkung: Zu einer Faktorisierung der Funktion $u = w^T v$ ergeben sich die Teilfaktoren w_2 und v_1 als eindeutig, vergleiche dazu mit den Sätzen 3.15 und 3.19.

Lemma 3.12. Sei $T \in \mathbb{R}^{s \times s}$ regulär und für ein $i_0 \in \{1, \dots, s\}$ gelte $T_{i,i_0} \geq 0$ für $i = 1, \dots, s$.

Dann besitzt T (mindestens) eine der beiden folgenden Eigenschaften.

- a) Es gibt genau ein i mit $T_{i,i_0} > 0$ oder
- b) es gibt i, j mit $(T^{-1})_{i,j} < 0$.

Beweis: Seien o.B.d.A. $i_0 = 1$ und $I = \{i : T_{i,1} > 0\}$. Gilt Eigenschaft b) nicht (das heißt $(T^{-1})_{i,j} \geq 0, \forall i, j$), so folgt daraus

$$(T^{-1})_{j,i} = 0, \quad \forall j = 2, \dots, s, \quad i \in I,$$

und es gilt

$$\text{rank}((T^{-1})_{1:s,I}) = 1.$$

Da T regulär ist, ergibt dies wiederum, dass $|I| = 1$ sein muss, also besitzt T die Eigenschaft a), sofern Eigenschaft b) auf T nicht zutrifft. \square

Damit lässt sich eine erste Abschwächung der Voraussetzungen aus Satz 3.8 für einen der Faktoren formulieren.

Satz 3.13. Es sei die Faktorisierungsvoraussetzung 3.1 gegeben. Weiter sei (w, v) eine nichtnegative Faktorisierung von u und es gelte eine der folgenden Bedingungen.

1. Es gibt $\bar{x}, \bar{y}, \tilde{y}, \epsilon, \varepsilon \in \mathbb{R}_+^s$ mit $\epsilon_i, \varepsilon_i > 0, i = 1, \dots, s$, und $j_0 \in \{1, \dots, s\}$, so dass

$$w_i(\bar{x}_j) = \epsilon_i \delta_{i,j}, \quad \forall i, j \in \{1, \dots, s\}, \quad (3.6a)$$

$$v_i(\bar{y}_j) = \varepsilon_i \delta_{i,j}, \quad \forall i, j \in \{1, \dots, s\}, \quad j \neq j_0, \quad (3.6b)$$

$$v_i(\tilde{y}_i) = 0, \quad v_{j_0}(\tilde{y}_i) > 0, \quad \forall i \in \{1, \dots, s\}, \quad i \neq j_0. \quad (3.6c)$$

2. Es gibt $\bar{x}, \bar{y}, \tilde{x}, \epsilon, \varepsilon \in \mathbb{R}_+^s$ mit $\epsilon_i, \varepsilon_i > 0$, $i = 1, \dots, s$, und $j_0 \in \{1, \dots, s\}$, so dass

$$\begin{aligned} w_i(\bar{x}_j) &= \epsilon_i \delta_{i,j}, & \forall i, j \in \{1, \dots, s\}, j \neq j_0, \\ v_i(\bar{y}_j) &= \varepsilon_i \delta_{i,j}, & \forall i, j \in \{1, \dots, s\}, \\ w_i(\tilde{x}_i) &= 0, \quad w_{j_0}(\tilde{x}_i) > 0, & \forall i \in \{1, \dots, s\}, i \neq j_0. \end{aligned}$$

Dann sind alle nichtnegativen Faktorisierungen von u äquivalent.

Beweis: Da die Voraussetzungen an die Faktoren komplementär sind, beschränkt sich der Beweis auf die 1. Bedingung. Angenommen es existieren zwei Faktorisierungen $(w(x), v(y))$, $(\tilde{w}(x), \tilde{v}(y))$ und die zugehörige Transformationsmatrix aus (3.3a) und (3.3b) ist keine permutierte Diagonalmatrix. Aus (3.6a) folgt, dass T^{-1} komponentenweise nichtnegativ sein muss. Soll T^{-1} keine permutierte Diagonalmatrix sein, muss T nach Lemma 3.7 negative Einträge besitzen, wobei dafür durch die Bedingung aus (3.6b) nur Einträge der Spalte j_0 in Frage kommen. Sei dazu i_0 so gewählt, dass $T_{i_0, j_0} < 0$ gilt. Aufgrund von (3.6c) muss es weiter j_1, j_2 mit $j_1 \neq j_2$ und $T_{i_0, j_1}, T_{i_0, j_2} > 0$ geben. Aus der Regularität von T und $T_{i,j} \geq 0$, $\forall i, \forall j \neq j_0$, (wegen (3.6b)) folgt, dass $\exists i_1 : i_1 \neq i_0$, $T_{i_1, j_1} > 0$ oder $T_{i_1, j_2} > 0$. Sei o.B.d.A $T_{i_1, j_1} > 0$. Da mit T_{i_0, j_1} und T_{i_1, j_1} mindestens zwei Elemente aus Spalte j_1 positiv sowie alle weiteren nichtnegativ sind, folgt nach Lemma 3.12, dass T^{-1} einen negativen Eintrag besitzen muss. Das widerspricht der Nichtnegativität von $\tilde{w}(x)$ in Verbindung mit der Voraussetzung aus (3.6a). \square

Zur Veranschaulichung der Aussage des Lemmas 3.13 ist in Beispiel 3.14 eine Funktion konstruiert, welche nur eine Äquivalenzklasse an Faktorisierungen besitzt.

Beispiel 3.14. (Eindeutige Faktorisierung) Zu $\alpha \in \mathbb{R}$ ist die Hilfsfunktion $\phi_\alpha : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ definiert als

$$\phi_\alpha(x) = \begin{cases} \sin^2(\pi(x - \alpha)) & \text{für } \alpha \leq x < \alpha + 1, \\ 0 & \text{sonst.} \end{cases}$$

Mit $s = 3$ sind dazu die Faktoren

$$\begin{aligned} w(x) &= (\phi_0(x), \phi_1(x), \phi_2(x))^T, \\ v(y) &= (\phi_0(y) + \phi_2(y), \phi_1(y) + \phi_3(y), \phi_2(y) + \phi_3(y))^T \end{aligned}$$

gewählt. Dann sind alle Faktorisierungen von

$$u(x, y) = w_1(x)v_1(y) + w_2(x)v_2(y) + w_3(x)v_3(y)$$

äquivalent und liegen in der Äquivalenzklasse $[(w, v)]_R$. Die Faktorisierung erfüllt dabei die erste Bedingung aus Satz 3.13.

Eine Auswertung von u an den Stützstellen $\bar{x} = (1/2, 3/2, 5/2)$, $\bar{y} = (1/2, 3/2, 5/2, 7/2)$ führt auf die Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Für die später untersuchte diskrete Problemstellung einer Matrixfaktorisierung werden die Zerlegungen ebenfalls in Äquivalenzklassen eingeteilt. Mit R aus Definition 4.13 besitzt A nur eine Äquivalenzklasse von Faktorisierungen, nämlich $[(I_3, A)]_R$.

Die in Satz 3.13 formulierten Bedingungen ließen sich auch weiter auf disjunkte Teilfaktorisierungen von $u(x, y)$ ausweiten. Mit *disjunkten Teilfaktorisierungen* ist gemeint, dass es eine disjunkte Zerlegung I_1, \dots, I_ℓ von $I = \{1, \dots, s\}$ derart gibt, dass für alle $j = 1, \dots, \ell$ für die Teilfunktionen $w_i, v_i, i \in I_j$ die Voraussetzungen von Satz 3.13 gelten und an den zugehörigen Stellen $\bar{x}, \bar{y}, \tilde{y}$ beziehungsweise $\bar{x}, \bar{y}, \tilde{x}$ die Funktionen $w_i, v_i, i \notin I$ verschwinden.

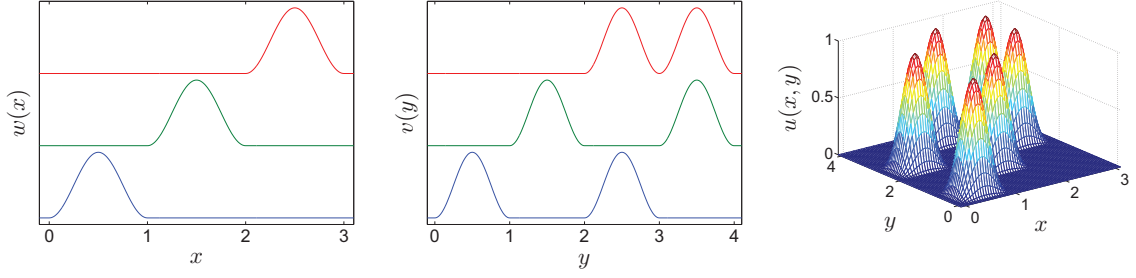


Abbildung 3.3: Dargestellt sind die Funktionen w , v und u aus Beispiel 3.14. Die hinreichenden Voraussetzungen für eine eindeutige Faktorisierung aus Satz 3.13 sind erfüllt. Alle Faktorisierungen liegen in derselben Äquivalenzklasse.

3.3 Einschränkungen der Menge aller Faktorisierungen

In den Sätzen 3.6 und 3.8 sind Bedingungen an die Faktoren formuliert, so dass alle möglichen Faktorisierungen in derselben Äquivalenzklasse liegen. Diese Bedingungen sind sehr einschränkend und nur in Sonderfällen erfüllt. Unabhängig davon lässt sich unter Gültigkeit bestimmter Voraussetzungen an die Ausgangsfaktoren die Vielfalt der möglichen Faktorisierungen einschränken.

In diesem Abschnitt werden Zusammenhänge unter den einzelnen Faktoren untersucht. Insbesondere werden Eingrenzungen der Menge aller Auflösungen analysiert. Anhand der Untersuchungen ist es möglich, Einschränkungen an einzelne Teilfaktoren anzugeben, welche durch zusätzliche Restriktionen (an die Faktoren) hervorgerufen werden.

3.3.1 Eindeutigkeit der Auflösung in einem Teilfaktor

Die hinreichenden Bedingungen für eine eindeutige Faktorisierung sind insbesondere dadurch sehr einschränkend, dass starke Forderungen an w und an v gestellt werden. Unter bestimmten Gegebenheiten führt jedoch bereits die Voraussetzung für nur einen Faktor auf eine Eindeutigkeit in einem Teilfaktor w_{i_0} oder v_{i_0} . Eine solche Einschränkung liefert der folgende Satz.

Satz 3.15. *Es sei die Faktorisierungsvoraussetzung 3.1 erfüllt. Weiter sei (w, v) eine nichtnegative Faktorisierung von u und es gebe $\bar{y} \in \mathbb{R}_+^s$, $\epsilon \in (\mathbb{R}_+ \setminus \{0\})^s$ sowie $i_0 \in \mathbb{N}$, $x_0 \in \mathbb{R}_+$ mit*

$$\begin{aligned} w_i(x_0) &= 0, & w_{i_0}(x_0) &> 0, & \forall i = 1, \dots, s, & i \neq i_0, \\ v_i(\bar{y}_j) &= \epsilon_i \delta_{i,j}, & & & \forall i, j = 1, \dots, s. \end{aligned}$$

Sei $(\tilde{w}(x), \tilde{v}(y))$ eine beliebige nichtnegative Faktorisierung von $u(x, y)$.

Dann gibt es ein j_0 und ein $\alpha > 0$, so dass gilt

$$\tilde{v}_{j_0}(y) = \alpha v_{i_0}(y).$$

Weiter gilt

$$\tilde{w}_j(x_0) = \alpha^{-1} \delta_{j_0,j} w_{i_0}(x_0), \quad \forall j = 1, \dots, s.$$

Beweis: *Es gelten mit einer regulären Matrix $T \in \mathbb{R}^{s \times s}$ die Beziehungen aus (3.3a) und (3.3b) und es gilt*

$$\begin{aligned} \tilde{w}_i(x_0) &= \sum_{j=1}^s w_j(x_0) (T^{-1})_{j,i} \geq 0, & i = 1, \dots, s, \\ \Rightarrow (T^{-1})_{i_0,i} &\geq 0, & i = 1, \dots, s. \end{aligned}$$

Nach Lemma 3.12 gilt für T eine der beiden folgenden Aussagen

- a) $\exists! j_0 : (T^{-1})_{i_0, j_0} > 0$ oder
- b) $\exists i_1, j_1 : T_{i_1, j_1} < 0$.

Angenommen T besitzt Eigenschaft b), dann folgt daraus

$$\tilde{v}_{i_1}(\bar{y}_{j_1}) = \sum_{l=1}^s T_{i_1, l} v_l(\bar{y}_{j_1}) = T_{i_1, j_1} v_{j_1}(\bar{y}_{j_1}) = T_{i_1, j_1} \epsilon_{j_1} < 0,$$

was der Nichtnegativität von \tilde{v} widerspricht. Damit kann Eigenschaft b) nicht gelten und es muss a) erfüllt sein. Sei gerade $\alpha^{-1} := (T^{-1})_{i_0, j_0} > 0$, wegen $(T^{-1})_{i_0, j} = 0$ für alle $j = 1, \dots, s$, $j \neq j_0$, hat $T_{j_0, \cdot}$ ebenfalls nur ein Element ungleich Null, nämlich $T_{j_0, i_0} = \alpha$, woraus folgt

$$\tilde{v}_{j_0}(y) = \sum_{i=1}^s T_{j_0, i} v_i(y) = \alpha v_{i_0}(y).$$

Für \tilde{w} gilt

$$\begin{aligned} \tilde{w}_j(x_0) &= \sum_{i=1}^s (T^{-1})_{i, j} w_i(x_0), & j &= 1, \dots, s, \\ &= (T^{-1})_{i_0, j} w_{i_0}(x_0), & j &= 1, \dots, s, \\ &= \alpha^{-1} \delta_{j_0, j} w_{i_0}(x_0), & j &= 1, \dots, s. \end{aligned}$$

□

Fazit 3.16. Erfüllen die Teilfaktoren aus $v(y)$ die Bedingungen (3.5b) und liegt zusätzlich der Teilfaktor $v_{i_0}(y)$ an einer Stelle x_0 isoliert vor, so ergibt sich die Funktion $v_{i_0}(y)$ abgesehen von Skalierung und Sortierung eindeutig in der Faktorisierung.

Beispiel 3.17. Die Faktorisierung der Funktion u aus Beispiel 3.11 ist nicht eindeutig. Nach Satz 3.15 findet sich jedoch der Teilfaktor $v_1(y)$ in jeder Faktorisierung wieder, vergleiche auch Abbildung 3.2. Für die Anwendung der Voraussetzungen aus Satz 3.15 lassen sich $\bar{y} = (\pi/2, \pi)^T$, $\epsilon = (1, 1)^T$, $i_0 = 1$, $x_0 = \pi/2$ wählen.

Bemerkung 3.18. Der Satz 3.15 ist die Verallgemeinerung von Satz 3.8. Werden die Voraussetzungen aus Satz 3.15 für alle $i_0 = 1, \dots, s$ (mit jeweils zugehörigen $x_0^{(i)} \in \mathbb{R}_+$) erfüllt, so liegt die Bedingung aus (3.5a) vor und die gesamte Faktorisierung ist eindeutig. Dabei folgt nach Satz 3.15 aus der Existenz der Stellen $x_0^{(i)}$ zunächst nur Eindeutigkeit in den Funktionen v_i , $i = 1, \dots, s$. Unmittelbar klar ist damit jedoch auch die Eindeutigkeit von w , vergleiche auch Lemma 3.32.

3.3.2 Einschränkungen an die Faktorisierung durch Eindeutigkeit oder Kenntnis einiger Teilfaktoren

Unter Umständen sind an die Faktorisierung der Funktion $u(x, y)$ zusätzliche Restriktionen durch spezielle Kenntnisse über einige der Teilfaktoren gestellt. Direkte Restriktionen an einige Teilfaktoren schränken indirekt auch andere Teilfaktoren ein. Gleiches gilt, falls einige Teilfaktoren in der Faktorisierung eindeutig sind. Derartige Zusammenhänge zwischen den Faktoren $w(x)$ und $v(y)$ und die Auswirkungen von Restriktionen oder der Eindeutigkeit einiger Teilfaktoren auf die gesamte Faktorisierung werden in diesem Teilabschnitt untersucht.

Aus der Kenntnis oder Eindeutigkeit von beispielsweise $w_i(x)$, $i = 1, \dots, s_0 < s$, lassen sich einerseits Einschränkungen für die komplementären Funktionsteile $v_j(y)$, $j = s_0 + 1, \dots, s$, sowie andererseits auch für die zugehörigen Funktionsteile $v_j(y)$, $j = 1, \dots, s_0$, schlussfolgern.

Zusammenhang zwischen komplementären Funktionenteilen

Inwiefern die komplementären Teilfaktoren $w_i(x)$, $i = 1, \dots, s_0 < s$, und $v_j(y)$, $j = s_0 + 1, \dots, s$, in Verbindung zueinander stehen, zeigt zunächst der Satz 3.19 sowie dessen Verallgemeinerung Korollar 3.21.

Satz 3.19. *Es seien die Faktorisierungsvoraussetzung 3.1 erfüllt und $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ so gewählt, dass $\text{rank}(U) = s$ mit $U_{i,j} = u(\bar{x}_i, \bar{y}_j)$, $i, j = 1, \dots, s$. Weiter bezeichne $X \in \mathbb{R}^{s \times s}$ die gesuchten Auswertungen $X_{i,j} = w_j(\bar{x}_i)$ und es sei $\tilde{U}\tilde{\Sigma}\tilde{V}^T = U$ eine Singulärwertzerlegung von U .*

Sind die Faktoren $w_i(x)$, $i = 1, \dots, s-1$, und somit auch $X_{:,1:s-1}$ bekannt, so ergibt sich mit der Bezeichnung aus 3.2 für $v_s(y)$ ein eindimensionaler Untervektorraum

$$v_s(y) \in \left\{ f(y) = \bar{t}^T \tilde{V}^T U^{-1} u_{\bar{x}}(y) : \bar{t} \in \mathbb{R}^s, \bar{t} \in \ker((\tilde{\Sigma}^{-1} \tilde{U}^T X_{:,1:s-1})^T) \right\}, \quad (3.7)$$

das heißt einzig die Skalierung ist unbekannt.

Beweis: Bezeichne $Y \in \mathbb{R}^{s \times s}$ die unbekannten Auswertungen $Y_{i,j} = v_i(\bar{y}_j)$. Desweiteren sind $\tilde{U}, \tilde{\Sigma}, \tilde{V} \in \mathbb{R}^{s \times s}$ und aus $\text{rank}(U) = s$ folgt, dass $\text{rank}(X) = \text{rank}(Y) = \text{rank}(\tilde{\Sigma}) = s$. Es gibt eine eindeutig bestimmte reguläre Matrix $T \in \mathbb{R}^{s \times s}$ mit

$$U = \underbrace{\tilde{U}\tilde{\Sigma}T}_{X} \underbrace{T^{-1}\tilde{V}^T}_{Y}, \quad (3.8)$$

wobei gilt

$$T_{:,1:s-1} = \tilde{\Sigma}^{-1} \tilde{U}^T X_{:,1:s-1}.$$

Der Teilfaktor $v_s(y)$ ergibt sich (vergleiche Lemma 3.3) mit einem $b \in \mathbb{R}^s$ als

$$v_s(y) = b^T u_{\bar{x}}(y).$$

Weiter gilt mit einem $\bar{t} \in \mathbb{R}^s$, $\bar{t} \in \ker(T_{:,1:s-1}^T) = \ker((\tilde{\Sigma}^{-1} \tilde{U}^T X_{:,1:s-1})^T)$,

$$Y_{s,:} = b^T U = \bar{t}^T \tilde{V}^T$$

und mit $b^T = \bar{t}^T \tilde{V}^T U^{-1}$ folgt

$$v_s(y) = \sum_{i=1}^s \left(\bar{t}^T \tilde{V}^T U^{-1} \right)_i u(\bar{x}_i, y).$$

□

Korollar 3.20. *Analog gilt Satz 3.19 auch für bekannte Faktoren $v_i(y)$, $i = 1, \dots, s-1$. Seien dazu die Matrizen $U, \tilde{U}, \tilde{\Sigma}, \tilde{V}$ definiert wie in Satz 3.19 und $Y_{i,j} = v_i(\bar{x}_j)$, $i = 1, \dots, s-1$, $j = 1, \dots, s$, bekannt.*

Dann gilt mit der Bezeichnung aus 3.2

$$w_s(x) \in \left\{ f(x) = (U^{-1} \tilde{U} \tilde{\Sigma} t)^T u_{\bar{y}}(x) : t \in \mathbb{R}^s, t \in \ker(Y_{1:s-1,:} \tilde{V}) \right\}.$$

Der Satz 3.19 zeigt, dass sich aus der Kenntnis von $w_i(x)$, $i = 1, \dots, s-1$, ein eindimensionaler Untervektorraum für $v_s(y)$ bestimmen lässt und lediglich die Skalierung von $v_s(y)$ unbekannt ist. Eine Verallgemeinerung dessen lässt sich ebenfalls angeben.

Korollar 3.21. *Es seien die Faktorisierungsvoraussetzung 3.1 erfüllt und $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ so gewählt, dass $\text{rank}(U) = s$ mit $U_{i,j} = u(\bar{x}_i, \bar{y}_j)$, $i, j = 1, \dots, s$. Weiter bezeichne $X \in \mathbb{R}^{s \times s}$ die gesuchten Auswertungen $X_{i,j} = w_j(\bar{x}_i)$ und sei $\tilde{U}\tilde{\Sigma}\tilde{V}^T = U$ eine Singulärwertzerlegung von U .*

Sind die Faktoren $w_i(x)$, $i = 1, \dots, s_0 < s$, und damit auch $X_{:,1:s_0}$ bekannt, so ergibt sich für $v_j(y)$, $j = s_0 + 1, \dots, s$, folgender Untervektorraum der Dimension $s - s_0$

$$v_j(y) \in \left\{ f(y) = \bar{t}^T \tilde{V}^T U^{-1} u_{\bar{x}}(y) : \bar{t} \in \mathbb{R}^s, \bar{t} \in \ker((\tilde{\Sigma}^{-1} \tilde{U}^T X_{:,1:s_0})^T) \right\}, \quad j = s_0 + 1, \dots, s.$$

Insbesondere gilt also für die gesuchten Funktionen

$$(v_{s_0+1}(y), \dots, v_s(y))^T = \bar{T}^T \tilde{V}^T U^{-1} u_{\bar{x}}(y)$$

mit einer Matrix $\bar{T} \in \mathbb{R}^{s \times (s-s_0)}$ vom Rang $s - s_0$ und

$$\bar{T} \in \ker((\tilde{\Sigma}^{-1} \tilde{U}^T X_{:,1:s_0})^T).$$

Bemerkung 3.22. *Analog zu Korollar 3.21 folgt aus bekannten Faktoren $v_i(y)$, $i = 1, \dots, s_0 < s$, mit den Bezeichnungen aus Korollar 3.20*

$$w_j(x) \in \left\{ f(x) = (U^{-1} \tilde{U} \tilde{\Sigma} t)^T u_{\bar{y}}(x) : t \in \mathbb{R}^s, t \in \ker(Y_{1:s-1,:} \tilde{V}) \right\}, \quad j = s_0 + 1, \dots, s.$$

Beispiel 3.23. *Das Beispiel 3.17 zeigt eine Anwendung von Satz 3.15. In der Faktorisierung der Funktion $u(x, y) = (w(x))^T v(y)$ aus Beispiel 3.11 ist der Faktor $v_1(y)$ eindeutig und kann somit als bekannt angesehen werden. Nach Korollar 3.20 ist auch $w_2(x)$ eindeutig, siehe dazu auch Abbildung 3.2.*

Zusammenhang zwischen zugeordneten Teilfunktionen

Zusätzlich zu den untersuchten Zusammenhängen unter den komplementären Teilfunktionen lassen sich aus der Kenntnis einer Funktion $w_{i_0}(x)$ auch bedingt Rückschlüsse auf den zugehörigen Teilfaktor $v_{i_0}(y)$ ziehen. Insgesamt ergibt sich ein affiner Raum der Dimension $s - 1$ für $v_{i_0}(y)$. Dazu ist der Satz 3.25 sowie vorbereitend das Lemma 3.24 als Vereinfachung der Sherman-Morrison-Woodbury Formel, [22], formuliert.

Lemma 3.24. *(Sherman-Morrison-Woodbury, siehe beispielsweise [22]) Seien $u, v \in \mathbb{R}^s$. Die Matrix $I_s - uv^T \in \mathbb{R}^{s \times s}$ ist genau dann regulär, wenn gilt $v^T u \neq 1$.*

Beweis: *Sei zunächst $I_s - uv^T$ regulär angenommen. Für den Fall $u = 0$ ist nichts zu zeigen, für $u \neq 0$ gilt*

$$(0, \dots, 0)^T \neq (I_s - uv^T)u = u - uv^T u = u(1 - v^T u) \Rightarrow 1 - v^T u \neq 0.$$

Sei weiter $(1 - v^T u) \neq 0$ und $I_s - uv^T$ nicht regulär angenommen. Dann muss $v \neq 0$ gelten und es gibt ein $x \in \mathbb{R}^s \setminus \{0\}$ mit

$$(0, \dots, 0)^T = (I_s - uv^T)x.$$

Damit gilt weiter $v^T x \neq 0$ und

$$0 = v^T (I_s - uv^T)x = v^T x - v^T uv^T x = (1 - v^T u)v^T x,$$

was weiter $1 - v^T u = 0$ folgert und einen Widerspruch zur Annahme bedeutet. Damit muss $I_s - uv^T$ regulär sein. \square

Satz 3.25. *Es seien die Faktorisierungsvoraussetzung 3.1 erfüllt und $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ so gewählt, dass $\text{rank}(U) = s$ mit $U_{i,j} = u(\bar{x}_i, \bar{y}_j)$, $i, j = 1, \dots, s$. Weiter bezeichne $X \in \mathbb{R}^{s \times s}$ die gesuchten Auswertungen $X_{i,j} = w_j(\bar{x}_i)$ und sei $\tilde{U}\tilde{\Sigma}\tilde{V}^T = U$ eine Singulärwertzerlegung von U .*

Ist der Faktor $w_{i_0}(x)$ und damit auch $X_{:,i_0}$ bekannt, so ergibt sich mit der Bezeichnung aus 3.2 und $t = \tilde{\Sigma}^{-1}\tilde{U}^T X_{:,i_0}$ für $v_{i_0}(y)$ der $(s-1)$ -dimensionale affine Raum $\mathcal{V}^{(i_0)}$,

$$v_{i_0}(y) \in \mathcal{V}^{(i_0)} = \left\{ f(y) = \bar{t}^T \tilde{V}^T U^{-1} u_{\bar{x}}(y) : \bar{t} \in \mathbb{R}^s, 0 = 1 - t^T \bar{t} \right\}. \quad (3.9)$$

Beweis: Sei $Y \in \mathbb{R}^{s \times s}$ definiert wie im Beweis von Satz 3.19 als $Y_{i,j} = v_i(\bar{y}_j)$. Mit $w_{i_0}(x)$ ist in Gleichung (3.8) die Spalte i_0 der Matrix T als

$$t = T_{:,i_0} = \tilde{\Sigma}^{-1} \tilde{U}^T X_{:,i_0}$$

bekannt. Weiter gilt für $\bar{t} \in \mathbb{R}^s$, $\bar{t}^T = (T^{-1})_{i_0,:}$

$$\begin{aligned} \text{rank}(U - X_{:,i_0} Y_{i_0,:}) < s &\Leftrightarrow \text{rank}(\tilde{U}\tilde{\Sigma}\tilde{V}^T - \tilde{U}\tilde{\Sigma}t\bar{t}^T\tilde{V}^T) < s \\ &\Leftrightarrow \text{rank}(I_s - t\bar{t}^T) < s \end{aligned}$$

und mit Hilfe von Lemma 3.24

$$\Leftrightarrow 0 = 1 - t\bar{t}^T.$$

Durch rückwärtige Substitution analog zum Beweis von Satz 3.19 folgt der Zusammenhang aus (3.9). \square

Die Idee aus Satz 3.25 lässt sich von $w_{i_0}(x)$ auf $v_{i_0}(y)$ übertragen.

Wie Satz 3.25 zeigt, ergibt sich bei Kenntnis einer Funktion $w_{i_0}(x)$ für die entsprechend zugehörige Funktion $v_{i_0}(y)$ ein affiner Unterraum der Dimension $s-1$. Dieser lässt sich unter Umständen noch weiter einschränken.

Korollar 3.26. *Es seien die Faktorisierungsvoraussetzung 3.1 erfüllt und $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ so gewählt, dass $\text{rank}(U) = s$ mit $U_{i,j} = u(\bar{x}_i, \bar{y}_j)$, $i, j = 1, \dots, s$. Weiter sei $\tilde{U}\tilde{\Sigma}\tilde{V}^T = U$ eine Singulärwertzerlegung von U .*

Ist zusätzlich eine Teilmenge $I \subset \{1, \dots, s\}$ bekannt, so dass für den Teilfaktor $v_{i_0}(y)$ gilt $v_{i_0}(\bar{y}_i) = 0$, $\forall i \in I$, so lässt sich für \bar{t} aus (3.9) die Einschränkung

$$v_{i_0}(y) \in \left\{ f(y) = \bar{t}^T \tilde{V}^T U^{-1} u_{\bar{x}}(y) : \bar{t} \in \mathbb{R}^s, \bar{t}^T \tilde{V}_{:,i}^T = 0, \forall i \in I \right\} \quad (3.10)$$

angeben.

Nach Korollar 3.21 lässt sich der Unterraum für v_{i_0} noch weiter einschränken. Seien dazu außer $w_{i_0}(y)$ weitere Teilfaktoren $w_j(x)$, $j \in J$, mit $J \subset \{1, \dots, s\} \setminus \{i_0\}$ bekannt. Daraus folgt die zusätzliche Einschränkung

$$v_{i_0}(y) \in \left\{ f(y) = \bar{t}^T \tilde{V}^T U^{-1} u_{\bar{x}}(y) : \bar{t} \in \mathbb{R}^s, \left(\tilde{\Sigma}^{-1} \tilde{U}^T X_{:,j} \right)^T \bar{t} = 0, \forall j \in J \right\}. \quad (3.11)$$

Bemerkung 3.27. *Die Bedingungen aus (3.10) und (3.11) sind unter Umständen untereinander abhängig. Für den Fall, dass sich die Funktionen $u(x, \bar{y}_i)$, $i \in I$, mit I aus Korollar 3.26 nur aus bekannten Faktoren $w_j(x)$, $j \in J$, mit J zu (3.11) zusammensetzen, also $v_\ell(\bar{y}_i) = 0$ für alle $i \in I$, $\ell \in \{1, \dots, s\} \setminus J$, gilt, folgt aus*

$$\left(\tilde{\Sigma}^{-1} \tilde{U}^T X_{:,j} \right)^T \bar{t} = 0, \quad \forall j \in J,$$

unmittelbar

$$\bar{t}^T \tilde{V}_{:,i}^T = 0, \quad \forall i \in I.$$

Obere und untere Einhüllende einzelner Funktionenteile

Weiterhin gelten natürlich die Nichtnegativitätsbedingungen

$$v_{i_0}(y) \in \{f(y) = \bar{t}^T u_{\bar{x}}(y) : \bar{t} \in \mathbb{R}^s, \bar{t}^T u_{\bar{x}}(y) \geq 0, \forall y \in \mathbb{R}_+\}.$$

Für die Systemrestriktion eines bekannten Teilfaktors $w_{i_0}(x)$ lässt sich zusätzlich eine obere Schranke für den zugeordneten Teilfaktor $v_{i_0}(y)$ angeben.

Korollar 3.28. *Es sei die Faktorisierungsvoraussetzung 3.1 erfüllt. Ist $w_{i_0}(x)$ bekannt, so gilt zusätzlich zu (3.9)*

$$v_{i_0}(y) \in \{f(y) = \bar{t}^T u_{\bar{x}}(y) : \bar{t} \in \mathbb{R}^s, \bar{t}^T u_{\bar{x}}(y) \leq \bar{v}_{i_0}(y), \forall y \in \mathbb{R}_+\}$$

mit der oberen Einhüllenden $\bar{v}_{i_0} : \mathbb{R}_+ \rightarrow \mathbb{R}_+$,

$$\bar{v}_{i_0}(y) = \inf_{\substack{x \in \mathbb{R}_+ \\ w_{i_0}(x) > 0}} \frac{u(x, y)}{w_{i_0}(x)},$$

wobei $0 \leq w_{i_0}(x)$ und $\exists x_0 \in \mathbb{R}_+ : w_{i_0}(x_0) > 0$ klar ist.

Beispiel 3.29. Erneut sei das Beispiel 3.11 mit den Faktoren $w(x)$ und $v(y)$ untersucht. Nach den Sätzen 3.15 und 3.19 ist der Faktor $w_2(x)$ bis auf Skalierung und Sortierung eindeutig. Ohne Beschränkung der Allgemeinheit seien die Skalierung identisch und die Sortierung $\hat{w}_2(x) = w_2(x)$. Nach Satz 3.25 gilt für jede nichtnegative Faktorisierung (\hat{w}, \hat{v}) von u , dass

$$\hat{v}_2(y) \in \mathcal{V}^{(i_2)} = \{\alpha \max(\sin(y), 0) + \max(\sin(y - \pi/2), 0) : \alpha \in \mathbb{R}\}.$$

Die Anwendung von Korollar 3.28 führt auf die obere Einhüllende $\bar{v}_2(y) = v_1(y) + v_2(y)$. Zusammen mit der Nichtnegativität von \hat{v}_2 lässt sich von dem affinen Raum $\mathcal{V}^{(i_2)}$ die Menge

$$\{\alpha \max(\sin(y), 0) + \max(\sin(y - \pi/2), 0) : \alpha \in [0, 1]\} \quad (3.12)$$

relevanter Lösungen $\hat{v}_2(y)$ abspalten. In Abbildung 3.4 sind ausgewählte Funktionen dieser Menge dargestellt.

Zusammenfassung aller Einschränkungen für zusammengehörige Teilfaktoren

Die aufgeführten Zusatzkenntnisse lassen sich zu Resultat 3.30 zusammenfassen.

Resultat 3.30. *Seien die Voraussetzungen aus Satz 3.25 sowie Korollar 3.26 erfüllt, die Bezeichnungen ebenfalls wie in 3.25 und 3.26 gewählt und J gegeben wie zu (3.11).*

Zusammenfassend ergibt sich für v_{i_0} die Darstellung

$$v_{i_0}(y) = \bar{t}^T \tilde{V}^T U^{-1} u_{\bar{x}}(y)$$

mit $\bar{t} \in \mathbb{R}^s$ und den Einschränkungen

$$0 = 1 - (\tilde{\Sigma}^{-1} \tilde{U}^T X_{:,i_0})^T \bar{t},$$

$$\bar{t}^T \tilde{V}_{i,:}^T = 0,$$

$$(\tilde{\Sigma}^{-1} \tilde{U}^T X_{:,j})^T \bar{t} = 0,$$

$$0 \leq \bar{t}^T \tilde{V}^T U^{-1} u_{\bar{x}}(y) \leq \inf_{x \in \mathbb{R}_+, w_{i_0}(x) > 0} \frac{u(x, y)}{w_{i_0}(x)},$$

$$\forall i \in I,$$

$$\forall j \in J,$$

$$\forall y \in \mathbb{R}_+.$$

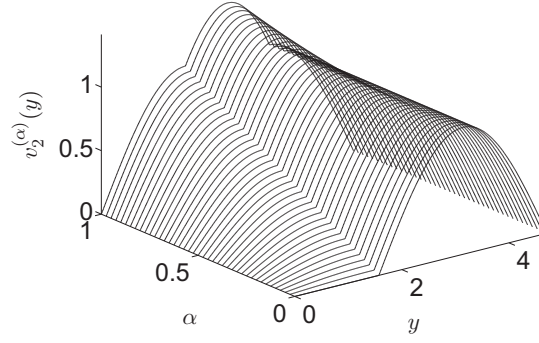


Abbildung 3.4: Dargestellt sind mögliche Teilfaktoren $\hat{v}_2(y) = v_2^{(\alpha)}(y)$ einer nichtnegativen Faktorisierung von u aus Beispiel 3.11 und Beispiel 3.29. Der Teilfaktor ist wie in (3.12) mit dem Parameter α konstruiert.

Ebenso lässt sich der umgekehrte Fall für den Faktor w zusammenfassen.

Resultat 3.31. *Es sei die Faktorisierungsvoraussetzung 3.1 erfüllt. Seien $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ so gewählt, dass $\text{rank}(U) = s$ mit $U_{i,j} = u(\bar{x}_i, \bar{y}_j)$, $i, j = 1, \dots, s$. Weiter bezeichne $Y \in \mathbb{R}^{s \times s}$ die gesuchten Auswertungen $Y_{i,j} = v_i(\bar{y}_j)$ und sei $U\tilde{\Sigma}V^T = U$ eine Singulärwertzerlegung von U . Dazu seien der Faktor $v_{i_0}(y)$ und damit auch $Y_{i_0,:}$ bekannt und es gebe weiter Teilmengen $I, J \subset \{1, \dots, s\}$, $i_0 \notin J$, so dass $w_{i_0}(\bar{x}_i) = 0$ für alle $i \in I$ gilt und $v_j(y)$ für $j \in J$ bekannt sind.*

Dann besitzt w_{i_0} folgende Darstellung

$$w_{i_0}(x) = (U^{-1}\tilde{U}\tilde{\Sigma}t)^T u_{\bar{y}}(x)$$

mit $t \in \mathbb{R}^s$ und den Einschränkungen

$$\begin{aligned} 0 &= 1 - Y_{i_0,:}\tilde{V}t, \\ \tilde{U}_{i,:}\tilde{\Sigma}t &= 0, & \forall i \in I, \\ Y_{j,:}\tilde{V}t &= 0, & \forall j \in J, \\ 0 \leq (U^{-1}\tilde{U}\tilde{\Sigma}t)^T u_{\bar{y}}(x) &\leq \inf_{y \in \mathbb{R}_+, v_{i_0}(y) > 0} \frac{u(x, y)}{v_{i_0}(y)}, & \forall x \in \mathbb{R}_+. \end{aligned}$$

3.3.3 Die Faktorisierung unter Kenntnis eines kompletten Blocks von Faktoren

Lösen lässt sich die Schwierigkeit der uneindeutigen Faktorisierung, indem das Problem durch Zusatzkenntnisse leicht abgewandelt wird. Ist neben $u(x, y)$ auch entweder $w(x)$ oder $v(y)$ bekannt, so lässt sich der fehlende Faktor trivialerweise eindeutig bestimmen, inklusive expliziter Kenntnis über Skalierung und Sortierung.

Lemma 3.32. *Es sei die Faktorisierungsvoraussetzung 3.1 erfüllt. Von der Zerlegung sei bereits der Teil $w \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s)$ beziehungsweise $v \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s)$ bekannt.*

Dann lässt sich dazu eindeutig $v(y)$ beziehungsweise $w(x)$ bestimmen.

Beweis: Seien zunächst der Fall des bekannten Faktors $w(x)$ angenommen und dazu $\bar{x}, \bar{y} \in \mathbb{R}_+^s$ so gewählt, dass $\det(U) \neq 0$ mit $U \in \mathbb{R}^{s \times s}$, $U_{i,j} = u(\bar{x}_i, \bar{y}_j)$. Weiter sei $X \in \mathbb{R}^{s \times s}$, $X_{i,j} = w_j(\bar{x}_i)$. Nach Lemma 3.3 lässt sich $v(y)$ mit einer regulären Matrix $T \in \mathbb{R}^{s \times s}$ konstruieren als

$$v(y) = Tu_{\bar{x}}(y).$$

Weiter gilt

$$u(\bar{x}_l, \bar{y}_\ell) = \sum_{i=1}^s w_i(\bar{x}_l) \sum_{j=1}^s T_{i,j} u(\bar{x}_j, \bar{y}_\ell), \quad l, \ell = 1, \dots, s,$$

$$U = XTU$$

und es ergibt sich $T = X^{-1}$ sowie $v(y)$ als

$$v(y) = X^{-1} u_{\bar{x}}(y). \quad (3.13)$$

Der Fall zur Bestimmung von $w(x)$ für eine bekannte Funktion $v(y)$ lässt sich analog zeigen. \square

Die Charakterisierung des zugehörigen inversen Problems

Bei einer geeigneten Wahl der Definitionsmengen D und \mathcal{Y} ist Eigenschaft 1 aus Definition 2.6 erfüllt. Mit Lemma 3.32 folgt die Forderung 2. Bedingung 3, die stetige Abhängigkeit der Lösung v von der rechten Seite (u, w) , zeigt der folgende Satz. Dabei sei beispielhaft nur der Fall eines bekannten Teilfaktors $w(x)$ zusätzlich zu $u(x, y)$ betrachtet.

Satz 3.33. *Es sei die Faktorisierungsvoraussetzung 3.1 erfüllt. Seien weiter $m_u, m_w, m_v > 0$, $\Omega_x = \Omega_y = \mathbb{R}_+$, $\Omega = \mathbb{R}_+^2$ und $\bar{x} \in \Omega_x$ fest gewählt sowie $D = \mathcal{X}^{(s)}$ und*

$$\begin{aligned} \mathcal{X}^{(s)} &= \{v \in \mathcal{C}(\Omega_y, \mathbb{R}_+^s) : \exists \bar{y} \in \Omega_y^s : \det((v_i(\bar{y}_j))_{i,j=1,\dots,s}) \neq 0, \|v_i\|_{\infty, \Omega_y} \leq m_v, \forall i = 1, \dots, s\}, \\ \mathcal{Y}^{(s)} &= \left\{ (u, w) \in \mathcal{C}(\Omega, \mathbb{R}_+) \times \mathcal{C}(\Omega_x, \mathbb{R}_+^s) : \begin{aligned} &\det((w_j(\bar{x}_i))_{i,j=1,\dots,s}) \neq 0, \\ &\|u\|_{\infty, \Omega} \leq m_u, \|w_i\|_{\infty, \Omega_x} \leq m_w, \forall i = 1, \dots, s, \\ &\exists v \in \mathcal{X}^{(s)} : u(x, y) = \sum_{i=1}^s w_i(x) v_i(y) \end{aligned} \right\}. \end{aligned}$$

Weiter seien die folgenden Metriken zu $\mathcal{X}^{(s)}$ und $\mathcal{Y}^{(s)}$ definiert

$$\begin{aligned} d_1(v, \tilde{v}) &= \max_{i=1,\dots,s} \|v_i - \tilde{v}_i\|_{\infty, \Omega_y}, \\ d_2((u, w), (\tilde{u}, \tilde{w})) &= \max \left\{ \|u - \tilde{u}\|_{\infty, \Omega}, \max_{i=1,\dots,s} \|w_i - \tilde{w}_i\|_{\infty, \Omega_x} \right\}. \end{aligned}$$

Dann gilt für alle $(u, w) \in \mathcal{Y}^{(s)}$: Für alle $\varepsilon > 0$ gibt es ein $\delta > 0$, so dass

$$\forall (\tilde{u}, \tilde{w}) \in \mathcal{Y}^{(s)} : d_2((u, w), (\tilde{u}, \tilde{w})) < \delta \Rightarrow d_1(v, \tilde{v}) < \varepsilon,$$

wobei für $v(y), \tilde{v}(y)$ gilt

$$u(x, y) = \sum_{i=1}^s w_i(x) v_i(y), \quad \tilde{u}(x, y) = \sum_{i=1}^s \tilde{w}_i(x) \tilde{v}_i(y).$$

Beweis: Seien $(u, w) \in \mathcal{Y}^{(s)}$ gewählt sowie dazu $X \in \mathbb{R}^{s \times s}$, $X_{i,j} = w_j(\bar{x}_i)$. Es ist X regulär. Weiter sei $\delta > 0$ so gewählt, dass $\forall (\tilde{u}, \tilde{w}) \in \mathcal{Y}^{(s)}$ mit $\tilde{X} \in \mathbb{R}^{s \times s}$, $\tilde{X}_{i,j} = \tilde{w}_j(\bar{x}_i)$ gilt

$$\|X - \tilde{X}\|_{\infty} < \delta \Rightarrow \|X^{-1}(\tilde{X} - X)\|_{\infty} < 1, \quad (3.14)$$

beispielsweise $\delta \leq s^{-3/2} \|X^{-1}\|_F^{-1}$. Es ist \tilde{X} regulär (durch die Wahl von $\mathcal{Y}^{(s)}$) und nach Satz 2.3.4 aus [22] gilt

$$\|\tilde{X}^{-1} - X^{-1}\|_{\infty} \leq \frac{\|\tilde{X} - X\|_{\infty} \|X^{-1}\|_{\infty}^2}{1 - \|X^{-1}(\tilde{X} - X)\|_{\infty}} =: \epsilon.$$

Weiter ergibt sich nach (3.13)

$$v(y) = X^{-1}u_{\bar{x}}(y), \quad \tilde{v}(y) = \tilde{X}^{-1}\tilde{u}_{\bar{x}}(y)$$

und es folgt mit $\max\{d_2((u, w), (\tilde{u}, \tilde{w})), \|X - \tilde{X}\|_\infty\} < \delta$ (wobei $\delta \leq s^{-3/2}\|X^{-1}\|_F^{-1}$ wegen (3.14))

$$\begin{aligned} d_1(v, \tilde{v}) &= \max_{i=1, \dots, s} \left\| \sum_{j=1}^s (X^{-1})_{i,j} u(\bar{x}_j, y) - (\tilde{X}^{-1})_{i,j} \tilde{u}(\bar{x}_j, y) \right\|_{\infty, \Omega_y} \\ &\leq \max_{i=1, \dots, s} \left\{ \sum_{j=1}^s \left\| (X^{-1})_{i,j} u(\bar{x}_j, y) - (\tilde{X}^{-1})_{i,j} \tilde{u}(\bar{x}_j, y) \right\|_{\infty, \Omega_y} \right\} \\ &\leq \max_{i=1, \dots, s} \left\{ \sum_{j=1}^s \left| (X^{-1})_{i,j} - (\tilde{X}^{-1})_{i,j} \right| \|u\|_{\infty, \Omega} + \left| (\tilde{X}^{-1})_{i,j} \right| \|u(\bar{x}_j, y) - \tilde{u}(\bar{x}_j, y)\|_{\infty, \Omega_y} \right\} \\ &< \max_{i=1, \dots, s} (\epsilon m_u + \|\tilde{X}^{-1}\|_\infty \delta) \leq \max_{i=1, \dots, s} (\epsilon m_u + (\|X^{-1}\|_\infty + \epsilon) \delta). \end{aligned}$$

Da für $\epsilon = \epsilon(\delta)$ gilt $\lim_{\delta \rightarrow 0} \epsilon(\delta) = 0$, folgt aus δ genügend klein die Behauptung. \square

Korollar 3.34. Für den Fall eines bekannten Faktorblocks $v(y)$ ergibt sich Eigenschaft 3 aus Definition 2.6 analog zu Satz 3.33.

Bei der Kenntnis eines Faktors: Ein korrekt gestelltes inverses Problem

Zusammenfassend ergibt sich daraus folgendes Fazit.

Fazit 3.35. Mit den Mengen $D = \mathcal{X}^{(s)}, \mathcal{Y}^{(s)}$, dem Vektor $\bar{x} \in \mathbb{R}^s$ und m_w aus Satz 3.33 sowie

$$\mathcal{X}_w^{(s)} = \{w \in \mathcal{C}(\Omega_x, \mathbb{R}_+^s) : \det((w_i(\bar{x}_j))_{i,j=1, \dots, s}) \neq 0, \|w_i\|_{\infty, \Omega_x} \leq m_w, \forall i = 1, \dots, s\}$$

und dem Operator

$$\begin{aligned} F_w : \mathcal{X}_w^{(s)} \times \mathcal{X}^{(s)} &\rightarrow \mathcal{Y}^{(s)}, \\ F_w(w, v) &= \left(\sum_{i=1}^s w_i(x) v_i(y), w \right) \end{aligned}$$

ergibt sich für die Bestimmung von $v(y)$ zu $u(x, y)$ und $w(x)$ mit den Aussagen aus 3.32 und 3.33 ein nach Definition 2.6 korrekt gestelltes inverses Problem.

Gleiches folgt für den Fall eines bekannten Faktorblocks $v(y)$ bei geeigneter Wahl des Operators, der Definitions- und Bildmengen sowie den zugehörigen Metriken.

Zusammenfassung

In diesem Kapitel wurde die kontinuierliche Faktorisierungsaufgabe 2.2 untersucht. Die Unbestimmtheit der Skalierung und der Umsortierung der Faktorenteile in w, v löste sich durch eine Einteilung aller Faktorisierungen in Äquivalenzklassen. Bezüglich der Eindeutigkeit der Lösung wurden das notwendige Kriterium 3.6 und das hinreichende Kriterium 3.8 formuliert, wobei bereits die notwendigen Voraussetzungen in aller Regel nicht erfüllt sind. Es bleibt jedoch festzuhalten, dass unter bestimmten Voraussetzungen einige Teilfaktoren eindeutig sind.

Weiter wurde untersucht, wie sich mittels bekannter Restriktionen für einzelne Teilfaktoren auch auf andere Teilfunktionen schließen lässt und wie stark solche zusätzlichen Restriktionen die Menge aller möglichen Faktorisierungen einschränken. Für einzelne Teilfaktoren ergeben sich dabei zunächst ein- oder mehrdimensionale Untervektorräume sowie affine Räume. Durch die Nichtnegativitätseigenschaft der Faktoren werden diese zu Mengen beschränkt.

In der Regel betreffen derartige Eingrenzungen aber nur wenige Teilfaktoren und sind insgesamt nicht so umfangreich, dass alle Faktorisierungen in einer Äquivalenzklasse liegen. Für die diskrete Faktorisierungsaufgabe der nichtnegativen Matrixfaktorisierung gilt dies analog, siehe im folgenden Kapitel den Abschnitt 4.4. Um aus der Menge aller Zerlegungen Faktorenpaare mit bestimmten Eigenschaften zu selektieren, wird die Aufgabenstellung zu der regularisierten Faktorisierungsaufgabe 2.5 modifiziert.

4 Die nichtnegative Matrixfaktorisierung

In Kapitel 3 wurde der kontinuierliche Produktansatz untersucht. Ausgangspunkt dieses Kapitels ist dessen diskrete Formulierung, einerseits ohne Störung als nichtnegative Matrixfaktorisierung sowie andererseits unter dem Einfluss von Störungen in Form einer nichtnegativen Niedrigrangapproximation.

Dazu werden zunächst einige grundlegende Eigenschaften zur Niedrigrangapproximation zusammengetragen. Weiter wird der später genutzte Ansatz zur Bestimmung einer nichtnegativen Matrixfaktorisierung vorgestellt. Ausgangspunkt der aus [56] abgeleiteten Idee ist eine abgeschnittene Singulärwertzerlegung der Matrix A . Aufbauend auf den Überlegungen aus Abschnitt 3.3 werden zum Abschluss für den theoretischen Fall des ungestörten Problems Zusammenhänge zwischen den einzelnen Teilfaktoren untersucht. Von Interesse ist dabei, welche indirekten Auswirkungen zusätzlich bekannte Restriktionen (an einige Teilfaktoren) auf andere Teilfaktoren und damit die Vielfalt aller möglichen Faktorisierungen haben. Denselben Effekt (wie zusätzliche Restriktionen in Form bekannter Teilfaktoren) haben Teilfaktoren, die in der Zerlegung eindeutig sind.

Zunächst wird ein kurzer Überblick zu Arbeiten bezüglich der nichtnegativen Matrixfaktorisierung gegeben. Dabei wird auch auf die in dieser Arbeit nicht weiter untersuchte Frage zur Existenz einer Zerlegung in nichtnegative Faktoren vollen Ranges eingegangen.

4.1 Grundlagen zur nichtnegativen Matrixfaktorisierung

Zum inversen Problem der nichtnegativen Matrixfaktorisierung stellen sich die zwei zentralen Fragen einerseits nach der Existenz einer Faktorisierung in nichtnegative Matrizen vollen Ranges sowie andererseits nach deren Eindeutigkeit. Wesentliche Resultate wurden dabei von Thomas ([84], 1974), Campbell und Poole ([8], 1981), Tam ([83], 1981), Chen ([10], 1984) sowie Paatero und Tapper ([67], 1994) erzielt.

Der ersten Frage nach der Existenz einer Aufspaltung in Faktoren vollen Ranges widmeten sich zunächst Thomas, Campbell und Poole sowie Tam und stellten diesbezüglich Kriterien an die Matrix A auf, siehe dazu [8, 83, 84]. Unabhängig davon waren es später Cohen und Rothblum, [13]: Als nichtnegativer Rang einer nichtnegativen Matrix $\text{rank}_+(A)$ wird die kleinstmögliche Anzahl an nichtnegativen dyadischen Produkten bezeichnet, welche addiert A ergeben; sie zeigten, dass im Fall $A \in \mathbb{R}_+^{k \times n}$ mit $\min\{k, n\} \leq 3$ gilt $\text{rank}_+(A) = \text{rank}(A)$ und gaben weiter für $n = k = 4$ ein Beispiel mit $\text{rank}_+(A) = \text{rank}(A) + 1$ an, welches sich ebenfalls in [84] findet (siehe auch Beispiel 4.22 dieser Arbeit).

Als Erweiterung von [8, 83, 84] war es Chen, der sich in [10] mit der zweiten Frage nach der Eindeutigkeit der Faktorisierung beschäftigte, wobei für $A \in \mathbb{R}^{k \times n}$ diesbezüglich ausschließlich der Fall $\text{rank}(A) = n \leq k$ betrachtet wurde. Angegeben sind dazu notwendige und hinreichende Kriterien für die Eindeutigkeit der Faktorisierung. Aufgegriffen wurde das Problem der uneindeutigen Auflösung ebenfalls von Paatero und Tapper in [67] insofern, dass das Zwischenschalten einer regulären Matrix T in $XY = XT^{-1}TY$ unter bestimmten Umständen ebenfalls eine nichtnegative Matrixfaktorisierung ergibt. Dabei unterschieden Paatero und Tapper zwischen Matrixfaktoren X, Y , die nur durch Diagonalmatrizen in nichtnegative Faktoren XT^{-1} und TY

transformiert werden können sowie allen anderen, ohne Rücksicht auf Permutationen in den Teilfaktoren und führten den Begriff *p-rotierbar* oder *positiv rotierbar* für letztgenannte ein, also Faktoren, welche sich auch mittels einer Nichtdiagonalmatrix T in weitere Faktorisierungen überführen lassen. Anders als bei Chen, [10], fehlt dabei eine tiefergehende Untersuchung von Matrixprodukten, welche sich nur durch permutierte Diagonalmatrizen in weitere nichtnegative Faktorisierungen transformieren lassen. Einzig die „Daumenregel“, dass eine signifikante Anzahl an Nullelementen in allen Spalten von X und allen Zeilen von Y auf eine eindeutige Auflösung führt, wurde aufgestellt (siehe Zitat auf Seite 16).

Für den Normalfall einer nicht eindeutig faktorisierbaren Matrix warfen Paatero und Tapper in [67] die sich daran anschließende dritte Frage zur nichtnegativen Matrixfaktorisierung auf: Wie lässt sich aus der Menge aller möglichen Auflösungen objektiv ein Element selektieren?

„HOW TO FAVOUR ONE SOLUTION AMONG A SET OF P-ROTATEABLE SOLUTIONS? AND HOW TO DO IT IN A REPEATABLE OBJECTIVE WAY?“, Paatero, Tapper, [67].

Zur Umsetzung einer Faktorenauswahl wurde in [67] die Möglichkeit zusätzlicher Restriktionen angeführt, wobei Paatero und Tapper zunächst den Einsatz anwendungsorientierter quadratischer Funktionen vorschlugen. Dies führt somit auf die Aufgabenstellung einer regularisierten nichtnegativen Matrixfaktorisierung.

Für klassische und neuere Zugänge bei der Berechnung von Matrixfaktorisierungen sei auf Kapitel 5 verwiesen, in welchem verschiedene Methoden vorgestellt werden.

Vor der Analyse der klassischen nichtnegativen Matrixfaktorisierung in Bezug auf Eindeutigkeit der Lösung werden einige wichtige Eigenschaften zur Niedrigrangapproximation einer Matrix zusammengetragen, welche im späteren Verlauf der Arbeit benötigt werden und auf welchen der in dieser Arbeit angewendete Ansatz zur Erzeugung einer nichtnegativen Matrixfaktorisierung basiert.

4.2 Niedrigrangapproximation einer Matrix

Die Niedrigrangapproximation einer Matrix A ist ein wichtiges Instrument der numerischen Mathematik. Dabei wird A durch eine Matrix B beziehungsweise dem Produkt von zwei Matrizen B, C mit im Allgemeinen $\text{rank}(B) \ll \text{rank}(A)$ beziehungsweise $\text{rank}(BC) \ll \text{rank}(A)$ angenähert.

4.2.1 Die Singulärwertzerlegung

Auf dem Gebiet der Niedrigrangapproximation spielt die Singulärwertzerlegung eine entscheidende Rolle. Von Eckard und Young 1936 erstmalig mit „*The approximation of one matrix by another of lower rank*“ [18] veröffentlicht, hat sie sich, auch bedingt durch die wachsenden Kapazitäten für Rechentechnik, zu einem wichtigen numerischen Werkzeug entwickelt.

Satz 4.1. (Singulärwertzerlegung, Eckard und Young, [18], Golub, [22]) Sei $A \in \mathbb{R}^{k \times n}$, dann existieren orthogonale Matrizen $U \in \mathbb{R}^{k \times k}$ und $V \in \mathbb{R}^{n \times n}$ sowie nichtnegative reelle Zahlen $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min\{k,n\}}$, Singulärwerte genannt, so dass gilt

$$A = U \Sigma V^T$$

mit einer aufgefüllten Diagonalmatrix

$$\Sigma \in \mathbb{R}^{k \times n}, \quad \Sigma_{i,j} = \begin{cases} \sigma_i & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

Die Spalten von $U = (u_1, \dots, u_k)$ und $V = (v_1, \dots, v_n)$ werden linksseitige und rechtsseitige Singulärvektoren genannt.

Für die Eindeutigkeit der Singulärwertzerlegung liefert folgendes Korollar eine Aussage.

Korollar 4.2. *Es seien $A \in \mathbb{R}^{k \times n}$ eine Matrix vom Rang s und $A = U\Sigma V^T$ eine Singulärwertzerlegung von A mit den paarweise verschiedenen Singulärwerten $\sigma_1 > \sigma_2 > \dots > \sigma_s$. Weiter sei $A = \tilde{U}\tilde{\Sigma}\tilde{V}^T$ ebenfalls eine Singulärwertzerlegung von A .*

Dann gilt $\Sigma = \tilde{\Sigma}$ und die ersten s Spalten von U und V sind bis auf das Vorzeichen eindeutig bestimmt

$$\begin{aligned} U_{:,i} &= \pm \tilde{U}_{:,i}, & i &= 1, \dots, s, \\ V_{:,i} &= \pm \tilde{V}_{:,i}, & i &= 1, \dots, s. \end{aligned}$$

Beweis: Sei o.B.d.A. $k \leq n$. Dass $\Sigma = \tilde{\Sigma}$ gilt ist klar, da in den Hauptdiagonalen jeweils die positiven Wurzeln der Eigenwerte von AA^T stehen, diese absteigend geordnet sind und die restlichen Einträge Null sind. Es ist $U_{:,i}$ ein Eigenvektor von AA^T zum Eigenwert σ_i^2 . Der Eigenraum zu σ_i^2 hat die Dimension eins, da alle Eigenwerte größer Null paarweise verschieden sind. Ebenso ist $\tilde{U}_{:,i}$ ein Eigenvektor zu σ_i^2 von AA^T , beide sind normiert und damit gilt $U_{:,i} = \pm \tilde{U}_{:,i}$. Analog folgt $V_{:,i} = \pm \tilde{V}_{:,i}$ über $A^T A$. \square

Für die weitere Verwendung werden als Abstandsmaße für Matrizen die Spektralnorm und die Frobeniusnorm definiert.

Definition 4.3. (Siehe beispielsweise [22]) Sei $A \in \mathbb{R}^{k \times n}$, dann wird

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^T A)}$$

als Spektralnorm bezeichnet und

$$\|A\|_F = \sqrt{\sum_{i=1}^k \sum_{j=1}^n A_{i,j}^2}$$

als Frobeniusnorm. Dabei ist $\lambda_{\max}(A^T A)$ der Betragswert des betragsmäßig größten Eigenwertes von $A^T A$.

Bemerkung 4.4. Die Frobeniusnorm und die Spektralnorm sind orthogonal invariant, das heißt für orthogonale Matrizen $U \in \mathbb{R}^{k \times k}$ und $V \in \mathbb{R}^{n \times n}$ gilt

$$\begin{aligned} \|A\|_2 &= \|UAV\|_2, \\ \|A\|_F &= \|UAV\|_F. \end{aligned}$$

4.2.2 Eigenschaften der Niedrigrangapproximation

Um den späteren Umgang mit der Niedrigrangapproximation zu vereinfachen seien folgende zwei Bezeichnungen eingeführt.

Bezeichnung 4.5. Zu einer Matrix $A \in \mathbb{R}^{k \times n}$ mit $A = (a_1, \dots, a_n)$, $a_i \in \mathbb{R}^k$ für $i = 1, \dots, n$, und einer Zahl $s \in \mathbb{N}$, $1 \leq s \leq n$, werden im weiteren Verlauf die Bezeichnungen

$$\begin{aligned} A_{:,1:s} &:= (a_1, a_2, \dots, a_s) \in \mathbb{R}^{k \times s}, \\ A_{:,1:s}^T &:= (A_{:,1:s})^T \in \mathbb{R}^{s \times k} \end{aligned}$$

verwendet.

Bezeichnung 4.6. Zu einer Matrix A sei mit Hilfe einer Singulärwertzerlegung $A = U\Sigma V^T$ die Niedrigrangapproximation $A^{(s)}$ definiert als

$$A^{(s)} = U\Sigma_{:,1:s}V_{:,1:s}^T.$$

Für die Niedrigrangapproximation $A^{(s)}$ gilt Folgendes.

Lemma 4.7. Seien $A \in \mathbb{R}^{k \times n}$ und $A = U\Sigma V^T$ eine Singulärwertzerlegung, dann gilt

$$\|A^{(s)} - A\|_2^2 = \sigma_{s+1}^2, \quad \|A^{(s)} - A\|_F^2 = \sum_{i=s+1}^{\min\{k,n\}} \sigma_i^2.$$

Beweis: Aufgrund der orthogonalen Invarianz der Spektral- und der der Frobeniusnorm gelten

$$\begin{aligned} \|U\Sigma_{:,1:s}V_{:,1:s}^T - A\|_2^2 &= \|\text{diag}(\sigma_1, \dots, \sigma_s, 0, \dots, 0) - \text{diag}(\sigma_1, \dots, \sigma_{\min\{k,n\}})\|_2^2 = \sigma_{s+1}^2, \\ \|U\Sigma_{:,1:s}V_{:,1:s}^T - A\|_F^2 &= \sum_{i=s+1}^{\min\{k,n\}} \sigma_i^2. \end{aligned}$$

□

Die additive Störung einer Matrix $A \in \mathbb{R}^{k \times n}$ mit $\text{rank}(A) = s < \min\{k, n\}$ führt oft auf eine Matrix $\tilde{A} = A + E$ mit $\text{rank}(A + E) = \tilde{s} > s$. Für eine Niedrigrangapproximation von \tilde{A} ist die Wahl des unbekannten Ranges s der störungsfreien Matrix A entscheidend. Diesbezüglich wird der ε -Rang einer Matrix als mögliches Entscheidungskriterium eingeführt.

Definition 4.8. Seien $A \in \mathbb{R}^{k \times n}$ und $\varepsilon > 0$. Dann ist der ε -Rang einer Matrix A definiert als

$$\text{rank}_\varepsilon(A) = \min_{B \in \mathbb{R}^{k \times n}, \|A - B\|_2 \leq \varepsilon} \text{rank}(B).$$

Korollar 4.9. Die Zahl $s = \text{rank}_\varepsilon(A)$ ist gleichbedeutend mit der Anzahl der Singulärwerte von A , die echt größer ε sind, also $\sigma_s > \varepsilon \geq \sigma_{s+1}$.

Beweis: Sei s beliebig mit $s < \text{rank}(A)$ (da $\varepsilon > 0$). Der Beweis erfolgt mit Hilfe der Aussage

$$\min_{B \in \mathbb{R}^{k \times n}, \text{rank}(B)=s} \|A - B\|_2 = \|A - A^{(s)}\|_2 = \sigma_{s+1}$$

aus Satz 4.10, vergleiche auch [22]. Demzufolge gilt

$$\min_{B \in \mathbb{R}^{k \times n}, \|A - B\|_2 \leq \varepsilon} \text{rank}(B) = \min_{i: \|A - A^{(i)}\|_2 \leq \varepsilon} i = \min_{i: \sigma_{i+1} \leq \varepsilon} i$$

und die Behauptung folgt direkt. □

Die Bestapproximation

Mit dem ε -Rang einer Matrix A und der Approximation durch $A^{(s)}$ stellt sich gleichzeitig die Frage nach einer Bestapproximation für eine Matrix. Der folgende Satz zeigt, dass es (im Sinne von $\|\cdot\|_2, \|\cdot\|_F$) unter allen Rang- s Matrizen keine bessere Approximation an A gibt als $A^{(s)}$.

Satz 4.10. (Golub, [22], Stewart, [80]) Sei $A \in \mathbb{R}^{k \times n}$ mit $\text{rank}(A) = r$ und den Singulärwerten $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$. Dann gilt für jede beliebige Matrix $B \in \mathbb{R}^{k \times n}$ mit $\text{rank}(B) = s \leq r$

$$\|A - B\|_2^2 \geq \sigma_{s+1}^2, \tag{4.1}$$

$$\|A - B\|_F^2 \geq \sum_{i=s+1}^r \sigma_i^2. \quad (4.2)$$

Wie in Lemma 4.7 gezeigt, ist beispielsweise $A^{(s)}$ eine Bestapproximation.

Beweis: Der Beweis für die Spektralnorm geht auf Golub und van Loan in [22] zurück; Stewart liefert in [80] eine Herleitung von Ungleichung (4.2). Beides wird nun folgend vorgestellt.

Für den Beweis seien o.B.d.A. $n \leq k$ und $U\Sigma V^T = A$ eine Singulärwertzerlegung von A sowie $V = [v_1, \dots, v_n]$. Sei B vom Rang s , so lässt sich eine Orthonormalbasis x_1, \dots, x_{n-s} für den Kern von B finden, $\ker(B) = \text{span}\{x_1, \dots, x_{n-s}\}$. Aufgrund der Dimensionen gibt es einen Vektor z mit

$$z \in \text{span}\{x_1, \dots, x_{n-s}\} \cap \{v_1, \dots, v_{s+1}\}$$

und $\|z\|_2^2 = 1$. Es folgt mit $Bz = 0$,

$$\begin{aligned} Az &= \sum_{i=1}^{s+1} \sigma_i (v_i^T z) u_i, \\ 1 = \|z\|_2^2 &= \left\| \sum_{j=1}^{s+1} \alpha_j v_j \right\|_2^2 = \sum_{j=1}^{s+1} \alpha_j^2 \|v_j\|_2^2 = \sum_{j=1}^{s+1} \alpha_j^2 \end{aligned}$$

die Abschätzung

$$\|A - B\|_2^2 \geq \|(A - B)z\|_2^2 = \|Az\|_2^2 = \sum_{i=1}^{s+1} \sigma_i^2 (v_i^T z)^2 = \sum_{i=1}^{s+1} \sigma_i^2 \alpha_i^2 \geq \sigma_{s+1}^2$$

aus (4.1).

In dem Beweis für die Ungleichung (4.2) sei B aus beliebigen orthonormalen Vektoren $x_i \in \mathbb{R}^k$, $i = 1, \dots, s$, gebildet, $B = \sum_{i=1}^s x_i y_i^T$. Nun genügt es zu zeigen, dass

$$\left\| A - \sum_{i=1}^s x_i y_i^T \right\|_F^2 \geq \|A\|_F^2 - \sum_{i=1}^s \sigma_i^2.$$

Dabei gilt

$$\begin{aligned} \left\| A - \sum_{i=1}^s x_i y_i^T \right\|_F^2 &= \text{tr} \left(\left(A - \sum_{i=1}^s x_i y_i^T \right)^T \left(A - \sum_{i=1}^s x_i y_i^T \right) \right) \\ &= \text{tr} \left(A^T A + \sum_{i=1}^s (y_i - A^T x_i)(y_i - A^T x_i)^T - \sum_{i=1}^s A^T x_i x_i^T A \right). \end{aligned}$$

Da $\text{tr}((y_i - A^T x_i)(y_i - A^T x_i)^T) \geq 0$ und $\text{tr}(A^T x_i x_i^T A) = \|Ax_i\|_F^2$ gelten, ist der Beweis vollständig, falls

$$\sum_{i=1}^s \|Ax_i\|_F^2 \leq \sum_{i=1}^s \sigma_i^2$$

gilt. Seien dazu $V = (V_1, V_2)$, wobei V_1 die ersten s Spalten von V umfasst, und $\Sigma = (\Sigma_1, \Sigma_2)$ eine analoge Zerlegung ist. Dann gilt

$$\begin{aligned} \|Ax_i\|_F^2 &= \sigma_s^2 + (\|\Sigma_1 V_1^T x_i\|_F^2 - \sigma_s^2 \|V_1^T x_i\|_F^2) \\ &\quad - (\sigma_s^2 \|V_2^T x_i\|_F^2 - \|\Sigma_2 V_2^T x_i\|_F^2) \\ &\quad - \sigma_s^2 (1 - \|V^T x_i\|_F^2). \end{aligned}$$

Dabei sind die letzten beiden Terme nichtnegativ beziehungsweise Null und es ergibt sich

$$\begin{aligned}
\sum_{i=1}^s \|Ax_i\|_F^2 &\leq s\sigma_s^2 + \sum_{i=1}^s (\|\Sigma_1 V_1^T x_i\|_F^2 - \sigma_s^2 \|V_1^T x_i\|_F^2) \\
&= s\sigma_s^2 + \sum_{i=1}^s \sum_{j=1}^s (\sigma_j^2 - \sigma_s^2) |v_j^T x_i|^2 \\
&= \sum_{i=1}^s \left(\sigma_s^2 + (\sigma_j^2 - \sigma_s^2) \sum_{j=1}^s |v_j^T x_i|^2 \right) \\
&\leq \sum_{i=1}^s (\sigma_s^2 + (\sigma_i^2 - \sigma_s^2)) \\
&= \sum_{i=1}^s \sigma_i^2.
\end{aligned}$$

□

Diese Grundlagen bezüglich der Singulärwertzerlegung werden nun dafür verwendet, die nichtnegative Matrixfaktorisierung aus Problemstellung 2.3 beziehungsweise die Aufgabe der nichtnegativen Niedrigrangapproximation aus 2.4 zu bearbeiten.

4.3 Matrixfaktorisierung mittels Transformation der Singulärvektoren

Die im vorherigen Abschnitt angeführten Überlegungen, speziell Satz 4.10, zeigen, dass sich die Probleme der Matrixfaktorisierung und der Niedrigrangapproximation in den Dimensionen deutlich reduzieren lassen. Zur Erinnerung: Aufgabe ist es, eine Faktorisierung beziehungsweise Niedrigrangapproximation der Matrix $A \in \mathbb{R}^{k \times n}$ mit nichtnegativen Faktoren $X \in \mathbb{R}^{k \times s}$, $Y \in \mathbb{R}^{s \times n}$ vom Rang s zu erzeugen.

Den Ansatz hierfür liefert eine Idee von Lawton und Sylvestre aus [56]. Bei einer Singulärwertzerlegung der Produktmatrix $A = XY$ ergeben sich für U und V die Strukturinformationen von X und Y . Demzufolge lassen sich die Faktoren X und Y auch rückwärtig aus den links- und rechtsseitigen Singulärvektoren konstruieren.

4.3.1 Das idealisierte Modell

Sei die störungsfreie Faktorisierungsaufgabe 2.3 betrachtet. Analog zu (3.8) für die kontinuierliche Faktorisierung liefert das multiplikative Einfügen einer geeigneten regulären Matrix $T \in \mathbb{R}^{s \times s}$ und deren Inversen in die Singulärwertzerlegung $U\Sigma V^T = A$ eine Transformation der abstrakten Faktoren der Singulärwertzerlegung

$$\begin{aligned}
A = U\Sigma V^T &= U\Sigma_{:,1:s} V_{:,1:s}^T \\
&= \underbrace{U\Sigma_{:,1:s} T^{-1}}_X \underbrace{TV_{:,1:s}^T}_Y \\
&= XY.
\end{aligned} \tag{4.3}$$

Ziel ist es dabei, die Matrix T so zu wählen, dass X und Y komponentenweise nichtnegativ sind.

4.3.2 Das korrigierte Modell

Für gestörte Daten ergibt sich bei vorgegebenen s in der Regel eine Niedrigrangapproximation der Matrix A vom Rang $s \ll \text{rank}(A)$. Um möglicherweise in den Singulärvektoren $U_{:,s+1:k}$ und $V_{:,s+1:n}^T$ enthaltene Informationen bezüglich X und Y nicht zu ignorieren, wird eine Matrix $T \in \mathbb{R}^{s \times z}$ ($s \leq z \ll \min\{k, n\}$) und deren Pseudoinverse T^+ (siehe beispielsweise [22]) verwendet

$$\begin{aligned} A &\approx \underbrace{U\Sigma_{:,1:z}}_X T^+ \underbrace{TV_{:,1:z}^T}_Y \\ &= XY. \end{aligned} \quad (4.4)$$

Bei stark gestörten Daten mit $z \gg s$ wird folgender allgemeiner Ansatz mit $S \in \mathbb{R}^{z \times s}$, $T \in \mathbb{R}^{s \times z}$ genutzt

$$\begin{aligned} A &\approx \underbrace{U\Sigma_{:,1:z}S}_X \underbrace{TV_{:,1:z}^T}_Y \\ &= XY. \end{aligned} \quad (4.5)$$

Bemerkung 4.11. Die Anwendung von $T \in \mathbb{R}^{s \times z}$ mit $z \geq s$ wie in (4.4) oder von S und T aus (4.5) liefert keine bessere Approximation an A bezüglich $\|\cdot\|_2$ oder $\|\cdot\|_F$ als die Verwendung einer regulären Matrix $T \in \mathbb{R}^{s \times s}$ wie in (4.3). In beiden Fällen liegt eine Niedrigrangapproximation vom Rang s oder kleiner vor, welche nach Satz 4.10 im Sinne der Spektralnorm oder der Frobeniusnorm nicht besser sein kann als $U\Sigma_{:,1:s}V_{:,1:s}^T$ aus (4.3).

4.3.3 Existenz und Eindeutigkeit der Transformation

Durch die Vorüberlegung aus den Gleichungen (4.3), (4.4) beziehungsweise (4.5) reduziert sich die Faktorisierungsaufgabe von X und Y auf die Bestimmung von T beziehungsweise T und S , also in der Dimension von $(n+k)s$ auf s^2 , sz beziehungsweise $2sz$. Dass dieses Vorgehen für störungsfreie Daten berechtigt ist, es also für existierende Faktoren X und Y eine eindeutig bestimmte Matrix T gibt, zeigt der folgende Satz, siehe [64].

Satz 4.12. (Neymeyr et al., [64]) Sei das störungsfreie Matrixprodukt $A = XY$ gegeben mit $X \in \mathbb{R}^{k \times s}$, $Y \in \mathbb{R}^{s \times n}$ und $\text{rank}(A) = s \leq \min\{k, n\}$. Dann gibt es eine eindeutig bestimmte Matrix $T \in \mathbb{R}^{s \times s}$, so dass sich X und Y aus der abgeschnittenen Singulärwertzerlegung von A wie in (4.3) ergeben.

Beweis: Die beiden linearen Abbildungen

$$\begin{aligned} \varphi_1 : \mathbb{R}^s &\rightarrow \mathbb{R}^k, & \varphi_1 : x &\mapsto U\Sigma_{:,1:s}x, \\ \varphi_2 : \mathbb{R}^s &\rightarrow \mathbb{R}^k, & \varphi_2 : x &\mapsto Xx \end{aligned}$$

haben das gleiche Bild, da $\text{rank}(X) = s$. Somit gibt es eine eindeutig bestimmte reguläre Matrix $T \in \mathbb{R}^{s \times s}$ mit

$$U\Sigma_{:,1:s} = XT, \quad U\Sigma_{:,1:s}T^{-1} = X.$$

Daraus ergibt sich in (4.3)

$$0 = U\Sigma_{:,1:s}(V_{:,1:s}^T - T^{-1}Y)$$

und da $U\Sigma_{:,1:s}$ den vollen Rang s mit $s \leq k$ hat und $A^{(s)} = A = XY$ gilt, folgt

$$V_{:,1:s}^T = T^{-1}Y, \quad TV_{:,1:s}^T = Y.$$

□

4.4 Nichteindeutigkeit der nichtnegativen Matrixfaktorisierung

Ein wichtiger Punkt bei der Bestimmung von Faktorisierungen oder Niedrigrangapproximationen ist die Frage nach deren Eindeutigkeit. Hierzu erfolgt eine Übertragung der Aussagen des Abschnitts 3.2 auf den Fall einer Matrixfaktorisierung.

Bezüglich Skalierung und Sortierung werden die nichtnegativen Faktorisierungen analog zu Kapitel 3 in Äquivalenzklassen eingeteilt. Sei dazu Folgendes definiert.

Definition 4.13. Gegeben seien die Dimensionen k, n sowie der Rang s . Dazu seien die Mengen

$$\begin{aligned}\mathcal{X} &= \{X \in \mathbb{R}_+^{k \times s} : \text{rank}(X) = s\}, \\ \mathcal{Y} &= \{Y \in \mathbb{R}_+^{s \times n} : \text{rank}(Y) = s\}, \\ \mathcal{T} &= \{T \in \mathbb{R}^{s \times s} : \text{rank}(T) = s \text{ und } T \text{ ist eine permutierte Diagonalmatrix}\}\end{aligned}$$

und die Äquivalenzrelation

$$\begin{aligned}R &\subset (\mathcal{X} \times \mathcal{Y}) \times (\mathcal{X} \times \mathcal{Y}), \\ (X, Y) \sim_R (\tilde{X}, \tilde{Y}) &:\Leftrightarrow \exists T \in \mathcal{T} : \tilde{X} = XT^{-1}, \tilde{Y} = TY\end{aligned}$$

definiert.

4.4.1 Notwendige und hinreichende Bedingungen

Die notwendigen und hinreichenden Voraussetzungen für eine eindeutige nichtnegative Matrixfaktorisierung übertragen sich von der kontinuierlichen Faktorisierungsaufgabe 2.2 auf die diskrete 2.3. Dazu wird folgende Faktorisierungsvoraussetzung eingeführt.

Faktorisierungsvoraussetzung 4.14. Sei $A \in \mathbb{R}_+^{k \times n}$ derart gegeben, dass $\text{rank}(A) = s$ mit $s \geq 2$ gilt und es $X \in \mathbb{R}_+^{k \times s}$, $Y \in \mathbb{R}_+^{s \times n}$ gibt mit $A = XY$.

Korollar 4.15. (Notwendige Bedingung) Es seien die Faktorisierungsvoraussetzung 4.14 erfüllt sowie A derart, dass alle nichtnegativen Faktorisierungen von A bezüglich R aus 4.13 in derselben Äquivalenzklasse liegen. Sei dies weiter $[(X, Y)]_R$ mit $X \in \mathcal{X}$, $Y \in \mathcal{Y}$ und $A = XY$.

Dann gilt

$$\begin{aligned}\forall i, j = 1, \dots, s, i \neq j, \exists l \in \{1, \dots, k\} : & \quad X_{l,i} > 0, \quad X_{l,j} = 0, \\ \forall i, j = 1, \dots, s, i \neq j, \exists l \in \{1, \dots, n\} : & \quad Y_{i,l} > 0, \quad Y_{j,l} = 0.\end{aligned}$$

Der Beweis erfolgt analog zu dem des Satzes 3.6.

Korollar 4.16. (Hinreichende Bedingung) Es seien die Faktorisierungsvoraussetzung 4.14 erfüllt und $A = XY$ eine Faktorisierung mit $X \in \mathbb{R}_+^{k \times s}$, $Y \in \mathbb{R}_+^{s \times n}$.

Gibt es weiter Vektoren $\xi \in \{1, \dots, k\}^s$, $\chi \in \{1, \dots, n\}^s$ und $\varepsilon, \epsilon \in (\mathbb{R}_+ \setminus \{0\})^s$, so dass

$$\begin{aligned}X_{\xi_i, j} &= \varepsilon_i \delta_{i, j}, & i, j = 1, \dots, s, \\ Y_{j, \chi_i} &= \varepsilon_i \delta_{i, j}, & i, j = 1, \dots, s,\end{aligned}$$

gilt, so sind alle nichtnegativen Faktorisierungen von A äquivalent.

Der Beweis erfolgt analog zu dem des Satzes 3.8.

Korollar 4.17. *Anderslautend bedeutet die Bedingung aus Korollar 4.16, dass es für $A \in \mathbb{R}_+^{k \times n}$ zwei Permutationsmatrizen $P \in \mathbb{R}^{k \times k}$, $Q \in \mathbb{R}^{n \times n}$, eine Diagonalmatrix vollen Ranges $D \in \mathbb{R}_+^{s \times s}$ sowie $\hat{X} \in \mathbb{R}_+^{(k-s) \times s}$ und $\hat{Y} \in \mathbb{R}_+^{s \times (n-s)}$ gibt mit*

$$PAQ = \begin{pmatrix} D & \hat{Y} \\ \hat{X}D & \hat{X}\hat{Y} \end{pmatrix}.$$

Es sind alle Faktorisierungen von PAQ äquivalent zu

$$\left(\begin{pmatrix} I_s \\ \hat{X} \end{pmatrix}, (D, \hat{Y}) \right)$$

und es gilt ohne Berücksichtigung der Skalierung für die ursprünglichen Faktoren $A = XY$

$$X = P^T \begin{pmatrix} I_s \\ \hat{X} \end{pmatrix}, \quad Y = (D, \hat{Y})Q^T.$$

Abschwächung der hinreichenden Bedingung

Die hinreichenden Voraussetzungen an die Faktoren für eine eindeutige Faktorisierung lassen sich geringfügig abschwächen. Die Voraussetzungen ergeben sich analog zu denen aus Satz 3.13 für den kontinuierlichen Produktansatz. Ein Beispiel dazu ist die Matrix

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

welche bereits in Beispiel 3.14 angeführt ist. Alle Faktorisierungen sind äquivalent zu (I_3, A) .

Eindeutigkeit in einem Teilfaktor

In Abschnitt 3.3.1 ist für den kontinuierlichen Produktansatz untersucht, unter welchen Voraussetzungen an die Faktoren einige Teilfaktoren in der Faktorisierung eindeutig sind. Dazu ist Satz 3.15 formuliert. Für den diskreten Ansatz gilt Folgendes.

Korollar 4.18. *Es seien die Faktorisierungsvoraussetzung 4.14 erfüllt und $A = XY$ eine Faktorisierung mit $X \in \mathbb{R}_+^{k \times s}$, $Y \in \mathbb{R}_+^{s \times n}$. Weiter gebe es Vektoren $\chi \in \{1, \dots, n\}^s$, $\epsilon \in (\mathbb{R}_+ \setminus \{0\})^s$, ein $\alpha > 0$ und Indizes $i_0 \in \{1, \dots, s\}$, $j_0 \in \{1, \dots, k\}$, so dass gilt*

$$\begin{aligned} X_{j_0,:} &= \alpha e_{i_0}^T, \\ Y_{j,\chi_i} &= \epsilon_i \delta_{i,j}, \quad i, j = 1, \dots, s. \end{aligned}$$

Dann ist der Teilfaktor $Y_{i_0,:}$ bis auf Skalierung eindeutig, das heißt zu jeder Faktorisierung $\tilde{X}\tilde{Y}$ gibt es i_1 und $\beta > 0$ mit

$$\tilde{Y}_{i_1,:} = \beta Y_{i_0,:}.$$

Der Beweis erfolgt analog zu dem des Satzes 3.15.

4.4.2 Der Spezialfall des vollen Ranges

In [10] ist zusätzlich eine notwendige Bedingung für eine eindeutige Faktorisierung an A für den Spezialfall $\text{rank}(A) = \min\{k, n\}$ gestellt. Diese ist in Lemma 4.19 formuliert und kann ebenfalls aus Korollar 4.15 abgeleitet werden. Dessen Ausweitung auf den allgemeinen Fall gilt trivialerweise nicht.

Lemma 4.19. *Sei $A \in \mathbb{R}^{k \times n}$ mit $k \geq n \geq 2$. Seien weiter $\text{rank}(A) = n$ und $A_{i_1, i_1} \geq \varepsilon A_{i_1, i_2}$, $i = 1, \dots, k$, mit einem $\varepsilon > 0$ und $i_1 \neq i_2$.*

Dann besitzt A keine eindeutige Faktorisierung.

Beweis: *Seien o.B.d.A. $i_1 = 1$, $i_2 = 2$ sowie $e_1 = (1, 0, \dots, 0)^T \in \mathbb{R}^s$, $e_2 = (0, 1, 0, \dots, 0)^T \in \mathbb{R}^s$. Dann sind durch*

$$A = AI_s = \underbrace{A(I_s - \varepsilon e_2 e_1^T)}_X \underbrace{(I_s + \varepsilon e_2 e_1^T)}_Y$$

zwei nichtnegative Faktorisierungen gegeben, welche nicht äquivalent sind.

Lemma 4.19 zeigt, dass, sofern A eindeutig faktorisierbar ist und den Rang n mit $n \geq 2$ besitzt, es eine Permutationsmatrix $P \in \mathbb{R}^{k \times k}$ gibt mit

$$(PA)_{i,i} = 0, \quad i = 1, \dots, n.$$

Die Ausweitung von Lemma 4.19 auf den Fall $\text{rank}(A) < \min\{k, n\}$ gilt selbstverständlich nicht, siehe Beispiel 4.20.

Beispiel 4.20. *Für folgende Matrix A sind alle Faktorisierungen äquivalent, vergleiche Korollar 4.16,*

$$A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$

4.4.3 Die Charakterisierung des Faktorisierungsproblems

An dieser Stelle sei eine Charakterisierung des diskreten Faktorisierungsproblems nach Hadamard aus Definition 2.6 angeführt. Zur allgemeinen Aufgabe der Matrixfaktorisierung wird zunächst ein bekanntes Resultat zur Frage nach der Existenz einer Zerlegung in Faktoren vollen Ranges angeführt.

Nichtnotwendige Existenz einer Zerlegung in Faktoren vollen Ranges

Eine Aussage zur Existenz einer nichtnegativen Matrixfaktorisierung liefert folgender Satz, vergleiche dazu auch Cohen und Rothblum mit [13].

Satz 4.21. *(Vergleiche beispielsweise Cohen, Rothblum, [13]) Für jede nichtnegative Matrix $A \in \mathbb{R}_+^{k \times n}$ mit $\text{rank}(A) = s$ existiert eine nichtnegative Matrixfaktorisierung $A = XY$ mit $X \in \mathbb{R}_+^{k \times s}$, $Y \in \mathbb{R}_+^{s \times n}$, falls gilt $\min\{k, n\} \leq 3$.*

Beweis: *Es sei o.B.d.A. $A \in \mathbb{R}_+^{k \times n}$ mit $k \geq n$. Der einzige nichttriviale Fall ist $n = 3$, $\text{rank}(A) = 2$ und $a_i \neq 0$, $i = 1, 2, 3$, mit $a_i \in \mathbb{R}^k$, $(a_1, a_2, a_3) = A$. Dann gibt es ein $\alpha \in \mathbb{R}^3 \setminus \{0\}$ mit*

$\alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_3 = 0$. Da $a_i \geq 0$ komponentenweise gilt, gibt es weiter i, j mit $\alpha_i \alpha_j < 0$. Sei o.B.d.A. $\alpha_1 > 0, \alpha_2 \geq 0, \alpha_3 < 0$, dann ist

$$A = (a_1, a_2) \begin{pmatrix} 1 & 0 & -\frac{\alpha_1}{\alpha_3} \\ 0 & 1 & -\frac{\alpha_2}{\alpha_3} \end{pmatrix} \quad (4.6)$$

eine nichtnegative Faktorisierung von A . □

Das folgende Beispiel aus [13, 84] verdeutlicht, dass die Ausdehnung auf $k, n > 3$ nicht gilt.

Beispiel 4.22. (Thomas, [84], Cohen, Rothblum, [13]) Sei

$$A = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix},$$

dann ist $\text{rank}(A) = 3$, aber es gibt keine Faktorisierung $A = XY$ mit $X \in \mathbb{R}_+^{4 \times 3}, Y \in \mathbb{R}_+^{3 \times 4}$.

Zur nichtnegativen Faktorisierung der ersten 3 Spalten von A werden nach Lemma 3.13 Vektoren der Form $(\alpha, \alpha, 0, 0)^T, (\beta, 0, \beta, 0)^T, (0, \gamma, 0, \gamma)^T$ mit $\alpha, \beta, \gamma > 0$ benötigt, aus welchen sich die vierte Spalte von A nicht mit ausschließlich nichtnegativen Faktoren konstruieren lässt.

Korollar 4.23. (Vergleiche ebenfalls Cohen, Rothblum, [13]) Aus Satz 4.21 lässt sich ableiten, dass jede nichtnegative Matrix $A \in \mathbb{R}_+^{k \times n}$ vom Rang zwei eine nichtnegative Matrixfaktorisierung besitzt. Insbesondere lässt sich eine solche Faktorisierung direkt berechnen.

Der Beweis und das Vorgehen zur Berechnung der Faktorisierung sind aus Übersichtlichkeitsgründen in Anhang D.1 verschoben.

Wie bereits in der Einführung erwähnt, sei durch den Hintergrund der Problemstellung aus Faktorisierungsaufgabe 2.2 und der daraus abgeleiteten nichtnegativen Matrixfaktorisierung 2.3 bei der Analyse der Faktorisierung nur der Fall einer, in Matrizen vollen Ranges s , faktorisierbaren Matrix untersucht.

Nichtnotwendige Eindeutigkeit einer Faktorisierung

Bei einer Charakterisierung des allgemeinen inversen Problems der nichtnegativen Matrixfaktorisierung mit Vollrangrestriktionen für die Faktoren nach Definition 2.6 ist bereits die erste Bedingung für ein „korrekt gestelltes“ Problem nicht erfüllt, sofern $\min\{n, k\} > 3$ gilt. Für die in dieser Arbeit analysierten Probleme ist jedoch speziell die zweite Eigenschaft von Interesse, die Eindeutigkeit der Faktorisierung. In Korollar 4.15 sind diesbezüglich notwendige Bedingungen an die Faktoren gestellt. In der Regel sind diese jedoch nicht erfüllt und es gibt keine eindeutige Faktorisierung.

Charakterisierung nach Hadamard: Ein schlecht gestelltes Problem

Damit lässt sich die Faktorisierungsaufgabe 2.3 auch in dem in dieser Arbeit untersuchten Anwendungskontext nach Hadamard als „schlecht gestellt“ einordnen. Selbst für den, nicht notwendig gegebenen, Fall einer, in nichtnegative Matrizen vollen Ranges, faktorisierbaren Matrix ist die Faktorisierung nur in Sonderfällen eindeutig und es handelt sich somit um ein schlecht gestelltes Problem.

4.5 Lösungseinschließungen unter zusätzlichen Restriktionen

Die bisherigen Untersuchungen in diesem Abschnitt basierten stets auf einem modellfreien Ansatz, das heißt für die Einschränkungen der Menge aller Faktorisierungsäquivalenzklassen ist nur die Nichtnegativität der Faktoren verantwortlich. Sind Teilfaktoren in den Spalten von X oder den Zeilen von Y bekannt (oder in der Faktorisierung eindeutig und somit bekannt), hat dies indirekt auch Auswirkungen auf andere Teilfaktoren. In Abschnitt 3.3 ist diesbezüglich der kontinuierliche Produktansatz untersucht. Auf den dortigen Ergebnissen aufbauend, wird nun für die diskrete Aufgabenstellung eine Übersicht erstellt, wie sich durch zusätzliche Restriktionen die Menge aller möglichen nichtnegativen Matrixfaktorisierungen einschränken lässt.

Seien dazu die Faktorisierungsvoraussetzung 4.14 gegeben, $U\Sigma V^T$ eine Singulärwertzerlegung von A und $X \in \mathbb{R}_+^{k \times s}$, $Y \in \mathbb{R}_+^{s \times n}$ die (gesuchten) Faktoren der Zerlegung.

1. Sind von dem Faktor X die ersten $s - 1$ Spalten $X_{:,1:s-1}$ bekannt, so ist (analog zu Satz 3.19) die Zeile $Y_{s,:}$ bis auf ihre Skalierung ebenfalls bekannt. Es gilt

$$Y_{s,:} \in \{t^T V_{:,1:s}^T : t \in \mathbb{R}^s, t \in \ker(((\Sigma_{:,1:s})^+ U^T X_{:,1:s-1})^T)\}.$$

2. Sind von dem Faktor X die ersten $s_0 < s$ Spalten bekannt, so ergibt sich (analog zu Korollar 3.21) für die Zeilen $Y_{j,:}$, $j = s_0 + 1, \dots, s$, ein $(s - s_0)$ -dimensionaler Untervektorraum

$$Y_{j,:} \in \{t^T V_{:,1:s}^T : t \in \mathbb{R}^s, t \in \ker(((\Sigma_{:,1:s})^+ U^T X_{:,1:s_0})^T)\}, \quad j = s_0 + 1, \dots, s.$$

3. Ist die Spalte $X_{:,i_0}$ bekannt, so ist die zugehörige Zeile $Y_{i_0,:}$ (analog zu Satz 3.25) Element des folgenden $(s - 1)$ -dimensionalen affinen Raums $\mathcal{Y}^{(i_0)}$

$$Y_{i_0,:} \in \mathcal{Y}^{(i_0)} = \{t^T V_{:,1:s}^T : t \in \mathbb{R}^s, 0 = 1 - t^T ((\Sigma_{:,1:s})^+ U^T X_{:,i_0})\}.$$

4. Ist für die Spalte $X_{:,i_0}$ eine Menge $I \subset \{1, \dots, k\}$ mit $X_{i,i_0} = 0$, $i \in I$, bekannt, so gilt (analog zu Korollar 3.26)

$$X_{:,i_0} \in \{U\Sigma_{:,1:s}\bar{t} : \bar{t} \in \mathbb{R}^s, \bar{t} \in \ker((U_{I,1:s}\Sigma_{1:s,1:s})^T)\}.$$

5. Ist $X_{:,i_0}$ bekannt, so gilt neben den Nichtnegativitätsrestriktionen (analog zu Satz 3.28)

$$Y_{i_0,:} \in \left\{ t^T V_{:,1:s}^T : t \in \mathbb{R}^s, 0 \leq t^T V_{i,1:s}^T \leq \min_{\substack{j=1,\dots,k, \\ X_{j,i_0} > 0}} \frac{A_{j,i}}{X_{j,i_0}}, i = 1, \dots, n \right\}.$$

6. Ist $X_{:,i_0}$ bekannt und gibt es zusätzlich zu den Indizes $i_1, \dots, i_s \in \{1, \dots, k\}$ einen Vektor $\varepsilon \in (\mathbb{R}_+ \setminus \{0\})^s$ mit

$$X_{i_\ell,j} = \varepsilon_j \delta_{j,\ell}, \quad j, \ell = 1, \dots, s,$$

so erzeugt das Anfügen von $X_{:,i_0}$ an A (aufbauend auf Satz 3.15) unskaliert Eindeutigkeit in der Zerlegung bezüglich $X_{:,i_0}$. Als neue Ausgangsmatrix ergibt sich dabei

$$A' = (A, X_{:,i_0}) \in \mathbb{R}^{k \times (n+1)}.$$

4.5.1 Einbinden mehrerer bekannter Teilfaktoren

Der Punkt 6 in der zuvor angeführten Auflistung zeigt eine Möglichkeit einzelne bekannte Teile von X oder Y in die Faktorisierung einzuarbeiten. Sind sowohl Teilfaktoren aus X und Y bekannt, lässt sich eine Erweiterung des multivariaten Modells wie folgt umsetzen.

Korollar 4.24. *Es seien die Faktorisierungsvoraussetzung 4.14 erfüllt und von den Faktoren die Teilkomponenten $X_{:,1:i_0}$ und $Y_{j_0+1:s,:}$ mit $1 \leq i_0 \leq j_0 < s$ bekannt. Die Erweiterung der Matrix A zu*

$$A' = \begin{pmatrix} A & X_{1:i_0,:} \\ Y_{:,j_0+1:s} & 0_{(s-j_0) \times i_0} \end{pmatrix} \in \mathbb{R}^{(k+s-j_0) \times (n+i_0)}$$

erzeugt nach Korollar 4.18 für den Fall, dass es zu einer Faktorisierung $A = XY$ einen Vektor $\varepsilon \in (\mathbb{R}_+ \setminus \{0\})^s$ sowie Indizes $i_1, \dots, i_s \in \{1, \dots, k\}$ mit

$$X_{i_\ell,j} = \varepsilon_j \delta_{j,\ell}, \quad j, \ell = 1, \dots, s,$$

gibt, unskaliert Eindeutigkeit in $X_{:,1:i_0}$. Analog folgt aus der Existenz von $\epsilon \in (\mathbb{R}_+ \setminus \{0\})^s$ sowie Indizes $i_1, \dots, i_s \in \{1, \dots, n\}$ mit

$$Y_{j,i_\ell} = \epsilon_j \delta_{j,\ell}, \quad j, \ell = 1, \dots, s,$$

bis auf Skalierung Eindeutigkeit in $Y_{j_0+1:s,:}$.

Bemerkung 4.25. *In Korollar 4.24 ist der Fall $i_0 > j_0$ uninteressant. Die Faktorenpaare (X_i, Y_i) , $i = j_0 + 1, \dots, i_0$, ließen sich jeweils direkt aus A extrahieren.*

Bemerkung 4.26. *Sei $A = XY$. Dann ist eine mögliche Faktorisierung $A' = X'Y'$ der Matrix aus Korollar 4.24 gegeben durch*

$$X' = \begin{pmatrix} X \\ 0_{(s-j_0) \times j_0} \end{pmatrix}, \quad Y' = (Y, \begin{pmatrix} I_{i_0} \\ 0_{(s-i_0) \times i_0} \end{pmatrix}). \quad (4.7)$$

4.5.2 Systemrestriktionen für spektroskopische Daten

Für eine zu faktorisierte Matrix von spektroskopischen Daten lassen sich auf verschiedene Weisen Teile des Faktors $\hat{A} = Y$ bestimmen. Einerseits können Zeilen von \hat{A} durch Einzelmessungen bekannt sein, worauf sich unter Umständen die Sätze 3.19 und 3.25 anwenden lassen, und andererseits bestehen in der Praxis Möglichkeiten, einzelne Zeilen von \hat{A} zu ermitteln.

Sind zwei Indizes $i \neq j$ und $\alpha \neq 1$ bekannt mit

$$\hat{C}_{j,:} = (\alpha \hat{C}_{i,1}, \hat{C}_{i,2}, \dots, \hat{C}_{i,s}), \quad (4.8)$$

so lässt sich

$$\hat{A}_{1,:} = \frac{1}{\hat{C}_{i,1}(1-\alpha)}(A_{i,:} - A_{j,:})$$

bestimmen. Analog ist es ebenfalls möglich, den absoluten Unterschied zweier Zeilen aus \hat{C} in lediglich einer Komponente ausnutzen. Seien dazu $i \neq j$, $\alpha \neq 0$ und

$$\hat{C}_{i,:} - \hat{C}_{j,:} = (\alpha, 0, \dots, 0), \quad (4.9)$$

dann ist $\hat{A}_{1,:}$ direkt berechenbar durch $\hat{A}_{1,:} = \alpha^{-1}(A_{i,:} - A_{j,:})$.

Anwenden lassen sich diese Szenarien durch einen Wiederholungsversuch mit einer geänderten Anfangskonzentration (Zusammenhang aus (4.8)) oder durch die Hinzugabe einer bestimmten Menge eines Stoffes zwischen den, im Vergleich zur Reaktionsdynamik, zeitnah beieinanderliegenden Messzeitpunkten t_i und t_{i+1} (Zusammenhang aus (4.9)). Das letztgenannte Vorgehen führt selbstverständlich nur auf eine Näherung.

Fazit

Die Schwierigkeiten der kontinuierlichen Faktorisierungsaufgabe 2.2 treten in gleicher Form bei der nichtnegativen Matrixfaktorisierung aus Faktorisierungsaufgabe 2.3 auf. In Bezug auf Eindeutigkeit der Auflösung können nur für Spezialfälle Ansprüche gestellt werden.

Durch spezielle Teilkenntnisse über die Faktoren X und Y lässt sich die Menge aller möglichen Faktorenpaare einschränken. Gleiches überträgt sich auf die Niedrigrangapproximation aus Faktorisierungsaufgabe 2.4. Trotz dessen ist die Auflösung in Äquivalenzklassen im Normalfall nicht eindeutig. Dies führt auf den Bedarf nach einer zusätzlichen Regularisierung der Faktoren; die Faktorisierungsaufgabe der regularisierten nichtnegativen Niedrigrangapproximation ist gestellt.

5 Klassische Zugänge zur Matrixfaktorisierung

In diesem Kapitel werden einige klassische Zugänge zur approximativen Matrixfaktorisierung vorgestellt und untersucht.

Zur Berechnung von Matrixfaktorisierungen sind in der Literatur eine Reihe von Methoden bekannt. Jeweils problemabhängig sind dabei verschiedene Restriktionen an die Faktoren gestellt und werden unterschiedliche Arten der Regularisierung angewendet. Bei der Hauptkomponentenanalyse wird eine Zerlegung mit unkorrelierten Daten in einem der Faktoren bestimmt. Ziel bei einer Unabhängigkeitsanalyse ist es, in einem Faktor möglichst unabhängige Daten zu erhalten. Die nichtnegative Matrixfaktorisierung erzeugt Faktoren mit nichtnegativen Einträgen, welche zusätzlichen Regularisierungen unterliegen können. Ein weiterer Algorithmus ist das kommerzielle Softwarepaket BTEM, welches speziell für die Anwendung auf spektroskopische Daten entwickelt wurde.

5.1 Hauptkomponenten- und Unabhängigkeitsanalyse

Die Hauptkomponenten- und die Unabhängigkeitsanalyse zerlegen einen großen Datensatz A in signifikante Faktoren X, Y jeweils durch eine Niedrigrangapproximation vorgegebenen Ranges s

$$A \approx XY, \quad A \in \mathbb{R}^{k \times n}, \quad X \in \mathbb{R}^{k \times s}, \quad Y \in \mathbb{R}^{s \times n}.$$

Idee dieser beiden Methoden ist es, diese Faktorisierung derart zu erzeugen, dass die Zeilen von Y beziehungsweise Spalten von X unkorreliert (Hauptkomponentenanalyse) beziehungsweise möglichst unabhängig untereinander (Unabhängigkeitsanalyse) sind. Auf mögliche Nichtnegativitätsrestriktionen von X oder Y wird keine Rücksicht genommen.

5.1.1 Die Hauptkomponentenanalyse (Principal Component Analysis)

Wie bereits eingangs erwähnt, wird bei der Hauptkomponentenanalyse eine Faktorisierung mit unkorrelierten Daten in einem Faktor bestimmt. Für Übersichtsarbeiten zur Hauptkomponentenanalyse ist unter anderem auf [9, 40, 43] verwiesen.

Varianz und Kovarianz

Für den späteren Umgang werden die Stichproben-Varianz, die Stichproben-Kovarianz sowie die Stichproben-Kovarianzmatrix benötigt. Details dazu lassen sich beispielsweise in [43] nachschlagen.

Zu zwei reellwertigen Zufallsgrößen Z_X und Z_Y sei die n -dimensionale Stichprobe (x_i, y_i) , $i = 1, \dots, n$, gegeben mit x_i , $i = 1, \dots, n$, unabhängig, (identisch verteilt) und die y_i ebenso. Mit der korrigierten Stichproben-Kovarianz

$$\text{Cov}(x, y) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y}) \quad (5.1)$$

ist ein erwartungstreuer Schätzer für die Kovarianz der Zufallsgrößen Z_X und Z_Y definiert. Dabei sind $\bar{x} = n^{-1} \sum x_i$ und $\bar{y} = n^{-1} \sum y_i$ Schätzer für die Erwartungswerte von Z_X und Z_Y . Weiter ist die korrigierte Stichproben-Varianz

$$\text{Var}(x) = \text{Cov}(x, x) = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (5.2)$$

ein erwartungstreuer Schätzer für die Varianz von Z_X .

Seien ferner Z_1, \dots, Z_k reellwertige Zufallsgrößen, zu welchen durch $A_{:,j} \in \mathbb{R}^k$, $j = 1, \dots, n$, eine n -dimensionale Stichprobe gegeben ist, zusammengefasst in $A \in \mathbb{R}^{k \times n}$. Für $i = 1, \dots, k$ seien $a_i = A_{i,:} \in \mathbb{R}^n$ die Stichprobeneinträge zu Z_i und diese jeweils unabhängig, (identisch verteilt). Dann ist $\Sigma \in \mathbb{R}^{k \times k}$,

$$\Sigma_{i,j} = \text{Cov}(a_i, a_j)$$

ein Schätzer für die Kovarianzmatrix der den a_i , $i = 1, \dots, k$, zugrunde liegenden Verteilungen.

Bemerkung 5.1. Für $x, y \in \mathbb{R}^k$ und $A \in \mathbb{R}^{k \times n}$ mit der zugehörigen Kovarianzmatrix Σ gilt

$$\text{Cov}(x^T A, y^T A) = x^T \Sigma y.$$

Die Hauptkomponenten und deren Gewichte

Die Hauptkomponenten $Y_1, \dots, Y_s \in \mathbb{R}^{1 \times n}$ werden als Linearkombinationen der Zeilen von A

$$Y_i = \tilde{x}_i^T A, \quad \tilde{x}_i \in \mathbb{R}^k,$$

derart bestimmt, dass sie untereinander unkorreliert sind und in ihrer Reihenfolge eine fallende Varianz besitzen.

Die Hauptkomponentenanalyse lässt sich für eine Rang- s Niedrigrangapproximation auf zwei Weisen anwenden.

1. Eine Hauptkomponentenanalyse von A liefert für

$$A = XY$$

unkorrelierte Hauptkomponenten in Y und deren Gewichte in X .

2. Andererseits liefert eine Hauptkomponentenanalyse von A^T die mit Y gewichteten unkorrelierten Hauptkomponenten X .

In diesem Abschnitt soll nur das Beispiel einer Hauptkomponentenanalyse von A betrachtet werden.

Bestimmung der ersten Hauptkomponente

Aufgabe ist es die erste Hauptkomponente Y_1 als Linearkombination

$$Y_1 = \tilde{x}_1^T A$$

so zu bestimmen, dass $\text{Var}(Y_1)$ maximal wird. Zur Beschränkung der Skalierung wird die Nebenbedingung $\tilde{x}_1^T \tilde{x}_1 = 1$ eingeführt und es ergibt sich

$$\tilde{x}_1 = \arg \max_{x \in \mathbb{R}^k, \|x\|_2=1} \text{Var}(x^T A) = \arg \max_{x \in \mathbb{R}^k, \|x\|_2=1} x^T \Sigma x,$$

woraus die Zielfunktion

$$f_1(x) = x^T \Sigma x - \lambda(x^T x - 1) \rightarrow \max$$

folgt. Der Gradient von $f_1(x)$

$$\begin{aligned} \nabla_x f_1(x) &= 2\Sigma x - 2\lambda x = 2(\Sigma - \lambda I_k)x, \\ \frac{\partial f_1}{\partial \lambda}(x) &= 1 - x^T x \end{aligned}$$

besitzt die Nullstellen

$$x = V_{:,i}, \quad \lambda = \sigma_i$$

mit den normierten Eigenvektoren $V_{:,i}$, $i = 1, \dots, k$, von Σ und den zugehörigen Eigenwerten σ_i . Unter diesen wird $f_1(x)$ für den Eigenvektor zum größten Eigenwert von A maximal und mit $\tilde{x}_1 = V_{:,1}$ gilt

$$\begin{aligned} Y_1 &= V_{:,1}^T A, \\ \text{Var}(Y_1) &= V_{:,1}^T \Sigma V_{:,1} = \sigma_1. \end{aligned}$$

Bestimmung der weiteren Hauptkomponenten

Alle weiteren Hauptkomponenten Y_i , $i = 2, \dots, s$, werden induktiv bestimmt. Dabei sollen diese unkorreliert mit den Vorherigen und von maximaler Varianz sein. Für die Zielfunktionen

$$\begin{aligned} f_i(x) &= x^T \Sigma x - \sum_{j=1}^{i-1} \lambda_j \text{Cov}(Y_{j,:}, x^T A) - \lambda_i(x^T x - 1), \quad i = 2, \dots, s, \\ &= x^T \Sigma x - \sum_{j=1}^{i-1} \lambda_j V_{:,j}^T \Sigma x - \lambda_i(x^T x - 1) \rightarrow \max \end{aligned}$$

ergeben sich

$$\begin{aligned} \nabla_x f_i(x) &= 2(\Sigma x - \lambda_i I_k)x - \sum_{j=1}^{i-1} \lambda_j V_{:,j}^T \Sigma, \\ \frac{\partial f_i}{\partial \lambda_j}(x) &= -V_{:,j}^T \Sigma x, \quad j = 1, \dots, i-1, \\ \frac{\partial f_i}{\partial \lambda_i}(x) &= 1 - x^T x, \end{aligned}$$

wobei als Lösungen

$$x = V_{:, \ell}, \quad \lambda_j = 0, \quad j = 1, \dots, i-1, \quad \lambda_i = \sigma_\ell$$

mit den normierten Eigenvektoren $V_{:, \ell}$, $\ell = i, \dots, k$, und den zugehörigen, der Größe nach absteigend geordneten, Eigenwerten σ_ℓ von Σ in Frage kommen. Damit ergeben sich zu $\tilde{x}_i = V_{:,i}$ die Hauptkomponenten

$$Y_i = V_{:,i}^T A, \quad i = 2, \dots, s,$$

und aus der Transformation $Y = \tilde{X} A$ mit $\tilde{X}^T = (\tilde{x}_1, \dots, \tilde{x}_s) = V_{:,1:s} \in \mathbb{R}^{k \times s}$ folgt über $(V_{:,1:s}^T)^+ = V_{:,1:s}$ eine Niedrigrangapproximation

$$A = VY = XY.$$

5.1.2 Die Unabhängigkeitsanalyse (Independent Component Analysis)

Einen ähnlichen Ansatz zur Hauptkomponentenanalyse, welche eine Faktorisierung in unkorrelierte Daten ergibt, stellt die Idee der Extraktion unabhängiger Komponenten aus den Zeilen der Datenmatrix dar. Die vorgestellten Ideen dieses Abschnittes zur Unabhängigkeitsanalyse sind hauptsächlich [14, 37–39] entnommen.

Motivation der Unabhängigkeitsanalyse

Seien Z_i , $i \in \mathbb{N}$, unabhängig, identisch verteilte und auf Erwartungswert $\mu = 0$ und Varianz $\sigma^2 = 1$ transformierte Zufallsgrößen. Dann konvergiert die Verteilungsfunktion von $S_n = n^{-1/2} \sum_{i=1}^n Z_n$ nach dem zentralen Grenzwertsatz, [74], für $n \rightarrow \infty$ gegen die Verteilungsfunktion der Standardnormalverteilung. In der Unabhängigkeitsanalyse wird daher das Ziel verfolgt, die einzelnen Komponenten möglichst stark abweichend von einer Normalverteilung zu konstruieren. Die Bestimmung der Unabhängigkeitskomponenten Y_1, \dots, Y_s als Linearkombinationen aus den Messdaten bleibt erhalten

$$Y = \tilde{X}A.$$

Abweichung von einer Normalverteilung

Eine Größe für die Abweichung einer Verteilung von einer Normalverteilung ist die Kurtosis

$$\text{kurt}(Y) = \mathbb{E}(Y^4) - 3(\mathbb{E}(Y^2))^2,$$

welche für eine normalverteilte Zufallsgröße Y den Wert 0 liefert und für fast alle anderen Verteilungen einen Wert ungleich Null, [38]. Zielstellung ist es daher, $|\text{kurt}(Y)|$ oder $(\text{kurt}(Y))^2$ zu maximieren. Zu beachten ist, dass die Kurtosis sensibel auf „Ausreißer“ reagiert und damit kein robustes Maß zur Bestimmung der Abweichung einer Zufallsverteilung von einer Normalverteilung ist, [38].

Einen Ausweg bietet die aus der Entropie abgeleitete Negentropie. Die Entropie geht auf die Arbeit „*A mathematical theory of communication*“ [76] von Shannon zurück und wird unter anderem in der Informationstechnologie eingesetzt. Interpretiert werden kann die Entropie als ein Zahlenwert für den Grad an Information, welche eine Beobachtung in sich birgt. Für eine diskrete Zufallsgröße z ist sie definiert als

$$H(z) = - \sum_i P(z = a_i) \ln P(z = a_i) \quad (5.3)$$

mit den möglichen Werten a_i für z . Bei einer stetigen Zufallsgröße z mit der Dichte p_z ergibt sich die Entropie als

$$H(z) = - \int p_z(\xi) \ln p_z(\xi) \, d\xi. \quad (5.4)$$

Für zufällige Beobachtungswerte liefert die Entropie einen großen Zahlenwert und für normalverteilte Zufallsgrößen den größten unter allen Zufallsgrößen mit gleicher Varianz, [16, 38, 69]. Daher ergibt sich mit der Negentropie

$$J(z) = H(\nu) - H(z)$$

und einer normalverteilten Zufallsgröße ν mit gleicher Varianz wie z eine nichtnegative Funktion $J(z)$, welche für eine normalverteilte Zufallsgröße den Wert Null liefert. Eine Approximation an die Negentropie liefert beispielsweise

$$J(z) \approx \frac{1}{12} \mathbb{E}(z^3)^2 + \frac{1}{48} \text{kurt}(z)^2,$$

welche wiederum das Problem der nicht robusten Kurtosis beinhaltet. Eine andere Approximation an $J(z)$ bietet

$$\tilde{J}(z) = (\mathbb{E}(G(z)) - \mathbb{E}(G(\nu)))^2$$

mit einer nichtquadratischen Funktion G . In dem FASTICA¹-Algorithmus von Hyvärinen, Gävert, Hurri und Särelä (vergleiche unter anderem [38]) kommen dafür beispielsweise

$$G(z) = G_1(z) = \frac{1}{a} \ln \cosh az$$

mit einer beliebigen Konstanten $1 \leq a \leq 2$ oder

$$G(z) = G_2(z) = -\exp\left(-\frac{z}{2}\right)$$

zum Einsatz.

Da bei der Maximierung von $\tilde{J}(Y) = \tilde{J}(\tilde{x}^T A)$ zufällige Startvektoren \tilde{x} gewählt werden, führen verschiedene Durchläufe des FASTICA-Algorithmus generell auch auf verschiedene Faktorisierungen in unabhängige Komponenten.

Zusammenfassung

Die Anwendung der Methoden PCA und ICA auf die Faktorisierungsaufgabe 2.3 ist nicht sinnvoll, da die Nichtnegativitätsrestriktionen nicht beachtet werden.

Die Methoden wären prinzipiell darauf erweiterbar, was jedoch nicht im Fokus dieser Arbeit steht. Weiter ist es in dem Anwendungskontext dieser Arbeit kritisch zu bewerten, die Faktoren sukzessiv einzeln zu bestimmen und das Faktorisierungsproblem nicht als zusammenhängende Aufgabe zu lösen.

5.2 Klassische Algorithmen zur nichtnegativen Matrixfaktorisierung

Zur Berechnung einer nichtnegativen Matrixfaktorisierung sind eine Reihe klassischer Algorithmen bekannt. Diese lassen sich grundsätzlich in drei allgemeine Klassen von Methoden unterteilen: *multiplikative Korrekturformeln*, *Gradientenverfahren* und *alternierende kleinste Quadrate Methoden*. In [3] sind eine Zusammenfassung der grundsätzlichen Methoden und Erläuterungen zu ausgesuchten speziellen Algorithmen zu finden; spezifische Arbeiten sind [12, 23, 47, 58–60, 75]. Zusätzlich sei auf Zdunek und Cichocki mit [87] verwiesen, die sich mit der Anwendung eines anderen Abstandsmaßes als $\|A - XY\|_F$ beschäftigten.

In diesem Abschnitt sollen mit den Algorithmen von Lee und Seung, [58], und Kim und Park, [47], zwei spezielle Ansätze vorgestellt werden.

¹MATLAB Code FastICA_25, bereitgestellt unter <http://www.cis.hut.fi/projects/ica/fastica/>.

Allgemeine Aufgabenstellung

Ziel ist es, zu einer nichtnegativen Matrix $A \in \mathbb{R}^{k \times n}$ eine nichtnegative Niedrigrangapproximation

$$\min_{X, Y \geq 0} f(X, Y), \quad (5.5)$$

$$f(X, Y) = \|A - XY\|_F^2 \quad (5.6)$$

mit $X \in \mathbb{R}^{k \times s}$, $Y \in \mathbb{R}^{s \times n}$, $s = \text{rank}_\varepsilon(A) \ll \min\{k, n\}$, für ein geeignetes $\varepsilon > 0$ zu bestimmen. Für den idealisierten Modellfall der nichtnegativen Matrixfaktorisierung 2.3 ist $\varepsilon < \sigma_r(A)$ zu wählen mit $\sigma_r(A)$ dem kleinsten nicht verschwindenden Singulärwert von A .

Eine Überführung des Problems in eine regularisierte nichtnegative Niedrigrangapproximation, ähnlich zu Faktorisierungsaufgabe 2.5, liefert die in der Literatur häufig auftretende Zielfunktion

$$f(X, Y) = \|A - XY\|_F^2 + \alpha \|X\|_F^2 + \beta \|Y\|_F^2, \quad \alpha, \beta \geq 0. \quad (5.7)$$

Bemerkung 5.2. Die Zielfunktionen aus (5.6) und (5.7) sind auf $\mathbb{R}_+^{k \times s} \times \mathbb{R}_+^{s \times n}$ nicht konvex. Somit sind die jeweils dazugehörigen Optimierungsprobleme ebenfalls nicht konvex.

Nachweis: Seien $e = (1, 0, \dots, 0)^T \in \mathbb{R}^k$, $\bar{e} = (1, 0, \dots, 0)^T \in \mathbb{R}^s$, $\tilde{e} = (1, 0, \dots, 0)^T \in \mathbb{R}^n$ und $\gamma \in \mathbb{R}_+ \setminus \{0\}$. Es ist mit f aus (5.7)

$$\begin{aligned} 0.5(f(\gamma e \bar{e}^T, 0) + f(0, \gamma \tilde{e} \tilde{e}^T)) &= \|A\|_F^2 + 0.5(\alpha + \beta)\gamma^2, \\ f(0.5\gamma e \bar{e}^T, 0.5\gamma \tilde{e} \tilde{e}^T) &= \|A\|_F^2 - A_{1,1}^2 + (A_{1,1} - 0.25\gamma^2)^2 + 0.25(\alpha + \beta)\gamma^2. \end{aligned}$$

Für eine konvexe Funktion folgt daraus die Forderung

$$\gamma^2 \leq 4(\alpha + \beta) + 8A_{1,1},$$

welche für feste Werte $A_{1,1}, \alpha, \beta \geq 0$ nicht für alle $\gamma \in \mathbb{R}_+$ zu halten ist.

Zu dem notwendigen Kriterium erster Ordnung einer lokalen Minimalstelle von f aus (5.6) unter Nichtnegativitätsrestriktionen sei Folgendes bemerkt. Notwendige beziehungsweise hinreichende Kriterien zweiter Ordnung sind in [65] zu finden.

Bemerkung 5.3. (Notwendiges Kriterium erster Ordnung) Die partiellen Ableitungen von f aus (5.6) sind

$$\frac{\partial f}{\partial X_{i,j}} = -2((A - XY)Y^T)_{i,j}, \quad \frac{\partial f}{\partial Y_{i,j}} = -2(X^T(A - XY))_{i,j}.$$

Ein lokales Minimum der restringierten Optimierungsaufgabe aus (5.5) erfüllt die Karush-Kuhn-Tucker Bedingungen, [49, 65]. Diese lauten für das spezielle Problem mit f aus (5.6)

$$\begin{aligned} X, Y &\geq 0, \\ (XY - A)Y^T, X^T(XY - A) &\geq 0, \\ X_{i,j}((XY - A)Y^T)_{i,j} &= 0, \quad \forall i = 1, \dots, n, j = 1, \dots, s, \\ Y_{i,j}(X^T(XY - A))_{i,j} &= 0, \quad \forall i = 1, \dots, s, j = 1, \dots, n. \end{aligned}$$

Die partiellen Ableitungen der Zielfunktion mit Regularisierungen aus (5.7) sind

$$\frac{\partial f}{\partial X_{i,j}} = 2(\alpha X - (A - XY)Y^T)_{i,j}, \quad \frac{\partial f}{\partial Y_{i,j}} = 2(\beta Y - X^T(A - XY))_{i,j}$$

und die notwendigen lokalen Optimalitätsbedingungen erster Ordnung ergeben sich sehr ähnlich zu denen bezüglich f aus (5.6).

5.2.1 Multiplikative Korrekturterme

Ein verbreiteter und sehr einfacher Algorithmus zur Lösung des nichtregularisierten Problems aus (5.5) ist durch die multiplikativen Korrekturformeln

$$\begin{aligned} Y_{i,j} &\leftarrow Y_{i,j} \frac{(X^T A)_{i,j}}{(X^T X Y)_{i,j}}, & i = 1, \dots, s, \quad j = 1, \dots, n, \\ X_{i,j} &\leftarrow X_{i,j} \frac{(A Y^T)_{i,j}}{(X Y Y^T)_{i,j}}, & i = 1, \dots, k, \quad j = 1, \dots, s, \end{aligned}$$

von Lee und Seung aus [58] gegeben. Weiter ist diesbezüglich in [58] gezeigt, dass für eine daraus konstruierte Folge von Iterierten $(X^{(l)}, Y^{(l)})$ für $l = 0, 1, 2, \dots$ die Auswertung von $\|A - X^{(l)} Y^{(l)}\|_F^2$ nicht steigend und nur für eine stationäre Iterierte invariant ist.

Einerseits ist der Algorithmus zwar sehr einfach, andererseits jedoch bergen die Korrekturterme auch einige Schwierigkeiten. Zunächst ist das Problem eines verschwindenden Nenners nicht behandelt. Weiter führen Startwerte mit $X_{i,j}^{(0)} = 0$ auf eine Folge von Iterierten $(X^{(l)}, Y^{(l)})$, $l = 1, 2, \dots$, mit $X_{i,j}^{(l)} = 0$, $\forall l$. Gonzalez in [23] und Lin in [60] zeigten weitere Schwächen des Algorithmus: So gilt beispielsweise unter bestimmten Voraussetzungen an A , $X^{(0)}$ und $Y^{(0)}$, dass

$$X_{i,j}^{(l)}, Y_{j,k}^{(l)} > 0, \quad \forall i, j, l, k.$$

Letztendlich sind von Gonzalez, [23], und Lin, [59], Verbesserungen des Algorithmus vorgeschlagen.

5.2.2 Alternierende kleinste Quadrate Methoden unter Nebenbedingungen

Die fehlende Konvexität der Optimierungsprobleme aus (5.6) und (5.7) birgt bei deren numerischer Lösung Schwierigkeiten. Einen Ausweg bietet folgender, unter anderen auf Karjalainen, [44], und Paatero, [67], zurückgehender, Ansatz. Die Optimierungsprobleme werden in jeweils zwei konvexe Teilprobleme aufgespalten, welche alternierend gelöst werden. Dazu wird einerseits für festes Y ein lineares Ausgleichsproblem unter Nebenbedingungen in X gelöst sowie andererseits zu festem X ein optimales Y ebenfalls als Lösung eines linearen Ausgleichsproblems unter Nichtnegativitätsrestriktionen bestimmt. Insgesamt ergeben sich aus den Optimierungsproblemen zu (5.6) und (5.7) jeweils die zwei Teilprobleme

$$\min_{X \geq 0} f_Y(X), \quad f_Y(X) = \|A - XY\|_F^2, \quad (5.8)$$

$$\min_{Y \geq 0} f_X(Y), \quad f_X(Y) = \|A - XY\|_F^2 \quad (5.9)$$

beziehungsweise

$$\min_{X \geq 0} f_Y(X), \quad f_Y(X) = \|A - XY\|_F^2 + \alpha \|X\|_F^2, \quad (5.10)$$

$$\min_{Y \geq 0} f_X(Y), \quad f_X(Y) = \|A - XY\|_F^2 + \beta \|Y\|_F^2. \quad (5.11)$$

Lemma 5.4. *Die Zielfunktionen aus (5.8) und (5.9) sind konvex, die der regularisierten Teilprobleme aus (5.10) und (5.11) sind für $\alpha, \beta > 0$ strikt konvex. Damit sind die Teilprobleme aus (5.8) bis (5.11) konvexe Optimierungsprobleme.*

Beweis: Der Nachweis wird für (5.9) und (5.11) geführt. Zunächst soll die Konvexität beziehungsweise strikte Konvexität der zugehörigen Zielfunktionen gezeigt werden. Eine Umformung

von $f_X(Y)$ aus (5.11) in eine eindimensionale Formulierung liefert mit $a \in \mathbb{R}^{kn}$, $\tilde{X} \in \mathbb{R}^{kn \times sn}$ und $y \in \mathbb{R}^{sn}$

$$\begin{aligned} a_{(j-1)k+i} &= A_{i,j}, \quad y_{(j-1)s+l} = Y_{l,j}, \quad i = 1, \dots, k, \quad j = 1, \dots, n, \quad l = 1, \dots, s, \\ \tilde{X}_{i+\nu k, l+\mu s} &= \begin{cases} 0 & \text{für } \nu \neq \mu, \\ X_{i,l} & \text{für } \nu = \mu \end{cases}, \quad i = 1, \dots, k, \quad l = 1, \dots, s, \quad \nu, \mu = 0, \dots, n-1, \end{aligned}$$

das quadratische Funktional

$$f_X(Y) = \|a - \tilde{X}y\|^2 + \beta\|y\|^2 = a^T a - 2a^T \tilde{X}y + y^T \underbrace{(\tilde{X}^T \tilde{X} + \beta I_{sn})}_{=:B} y. \quad (5.12)$$

Die Matrix B ist symmetrisch positiv semidefinit sowie für $\beta > 0$ positiv definit und damit ist $f_X(Y)$ konvex beziehungsweise für $\beta > 0$ strikt konvex. Der Beweis für (5.8) und (5.10) folgt analog. Da zusätzlich die Menge $Y = \mathbb{R}_+^{s \times n}$ konvex ist, sind die Optimierungsprobleme ebenfalls konvex. \square

Korollar 5.5. Die regularisierten Teilprobleme aus (5.10) und (5.11) besitzen für $\alpha, \beta > 0$ jeweils eindeutig bestimmte Lösungen.

Beweis: Für (5.11) gilt: Die Menge aller zulässigen Vektoren in (5.12) ist für y halboffen und an den Ebenen $y_i = 0$ abgeschlossen. Da die Zielfunktion zu dem Optimierungsproblem nach Lemma 5.4 für $\beta > 0$ strikt konvex ist und trivialerweise auch koerziv, besitzt die quadratische Funktion für $Y \in \mathbb{R}_+^{s \times n}$ ein eindeutig bestimmtes Minimum. Analoges gilt bei $\alpha > 0$ für das Optimierungsproblem aus (5.10). \square

Lösung der regularisierten Teilprobleme

Eine Herangehensweise, die Teilprobleme (5.8) und (5.9) zu lösen, liefern Kim und Park in [47] über die Aufspaltung in die n linearen Ausgleichsprobleme unter Nebenbedingungen

$$\min_{Y \geq 0} \|A - XY\|_F^2 \rightsquigarrow \min_{Y_{:,1} \geq 0} \|A_{:,1} - XY_{:,1}\|_2^2, \dots, \min_{Y_{:,n} \geq 0} \|A_{:,n} - XY_{:,n}\|_2^2,$$

welche direkt mit der *active-set*-Methode von Lawson und Hanson, [55], implementiert in der MATLAB Routine `lsqnonneg`, gelöst werden. Alternativ ergeben sich für die regularisierten Fälle die mit $\alpha, \beta > 0$ strikt konvexen Probleme

$$\min_{X \geq 0} \left\| \begin{pmatrix} A^T \\ 0_{s \times k} \end{pmatrix} - \begin{pmatrix} Y^T \\ \sqrt{\alpha} I_s \end{pmatrix} X^T \right\|_F^2, \quad \min_{Y \geq 0} \left\| \begin{pmatrix} A \\ 0_{s \times n} \end{pmatrix} - \begin{pmatrix} X \\ \sqrt{\beta} I_s \end{pmatrix} Y \right\|_F^2.$$

Nach der Wahl der Startmatrizen $X^{(0)}, Y^{(0)}$ ist damit die alternierende Lösung der genannten Teilprobleme die Iteration des Verfahrens. Die Dimensionen dieser Optimierungsstrategie sind demzufolge $\mathbb{R}^{k \times s}$ und $\mathbb{R}^{s \times n}$, was durchaus als umfangreich gewertet werden kann. Durch die Vorgabe von Unterräumen für X und Y , beispielsweise mittels einer abgeschnittenen Singularwertzerlegung von A , ließen sich die Probleme auf $\mathbb{R}^{s \times s}$ reduzieren und die Effizienz für $n, k \gg s$ deutlich steigern.

5.2.3 Anwendungsbereiche der nichtnegativen Matrixfaktorisierung in der Literatur

Multivariate Ansätze der Form $A = XY + E$ mit nichtnegativen Faktoren X, Y sind in der Literatur weit verbreitet. In dieser Arbeit ist der Fall einer Faktorisierung der Matrix A bei (relativ zum Produkt XY) kleinen Störungen untersucht.

Leicht abgewandelt davon sind Anwendungen in der Mustererkennung:

- Bild- und Gesichtserkennung, [26, 57],
- Schrift- beziehungsweise Ziffernerkennung, [25, 27],
- Klassifikation von Texten oder Objekten, [2, 3, 71].

Bei diesen Anwendungen ist im Normalfall für $s \ll \min\{k, n\}$ die Aufgabe einer Niedrigrangapproximation mit einer im Vergleich zum Produkt XY großen Störung E , also $\sigma_s(XY) \gg \sigma_1(E)$, gestellt. Weiter sind speziell in den Veröffentlichungen zur Bild- und Ziffernerkennung auch (zu $\|A - XY\|_F$) alternative Abstandsmaße angewendet, welche ebenfalls Änderungen in den Iterationsvorschriften nach sich ziehen.

5.3 Selbstmodellierende Zugänge in der Chemometrie

Ein weiterer klassischer Zugang zur Berechnung einer Niedrigrangapproximation einer Matrix A geht auf Lawton und Sylvestre: „*Self modelling curve resolution*“ [56] zurück. Der Ansatzpunkt für die darauf aufbauenden selbstmodellierenden Algorithmen ist eine Transformation der Faktoren der Singulärwertzerlegung von A . Letztendlich ist es diese Idee von Lawton und Sylvestre, welche die Basis der Überlegungen aus Kapitel 4 sowie der in Kapitel 6 vorgestellten neu entwickelten Methode PCD bildet.

Mit der Bezeichnung „*selbstmodellierender Algorithmus*“ wird eine Methode von „*modellbehafteten Algorithmen*“ abgehoben. Derart werden in der Chemometrie Verfahren bezeichnet, welche zusätzliche Restriktionen an eine Faktorisierung stellen, in Form eines kinetischen Modells für die Konzentrationsprofile des zugrunde liegenden Systems. Selbstmodellierende Algorithmen beschränken die Zerlegung lediglich auf nichtnegative Faktoren und begünstigen eine Faktorisierung mit typischen Einzelsubstanzspektren für einen dieser. Eine spezifiziertere Abgrenzung ist in [63] zu finden.

Notation

In diesem Abschnitt wird die Faktorisierungsaufgabe 2.5 (Bestimmung einer regularisierten nichtnegativen Niedrigrangapproximation) untersucht. Der üblichen Notation der Chemie folgend werden die Faktoren mit \hat{C} und \hat{A} bezeichnet, $\hat{C} = X$, $\hat{A} = Y$. Für Bezüge auf den idealisierten Produktansatz aus Faktorisierungsaufgabe 2.2 wird $t = x$ und $\hat{c}(t) = w(x)$ sowie $x = y$ und $\hat{a}(x) = v(y)$ verwendet.

5.3.1 Historische Entwicklung selbstmodellierender Algorithmen in der Spektroskopie

Lawton und Sylvestre bestimmten in dem 1971 publizierten Artikel „*Self modelling curve resolution*“ [56] die Teilfaktoren der Zeilen von \hat{A} als nichtnegative Linearkombinationen aus den Eigenvektoren von $A^T A$ zunächst für ein 2-Komponentensystem und ohne weitere Regularisierungen. Im Jahr 1974 folgte mit [82] eine publizierte Anwendung auf ein praktisches System, dem Multikomponentensystem

$$A + B \rightleftharpoons C,$$

inklusive eines Ausblicks auf die Anwendung in der Kinetik.

Später waren es Malinowski ([63], erste Auflage 1980), Kvalheim und Liang ([54], 1993) sowie Garland ([11, 68, 86, 88], ab 2000) die in der Weiterentwicklung des Ansatzes federführend waren.

Malinowski griff die Ideen von Lawton und Sylvestre in [63] auf und entwickelte diese weiter, Garland konzipierte darauf basierend das Softwarepaket BTEM.

Ausgehend von einer Singulärwertzerlegung wird eine Transformation der abstrakten Faktoren U und V^T in sinnvolle Profile berechnet. Entscheidender Punkt ist dabei die lokale Rang-1 Approximation der Matrix A durch dyadische Produkte zur einzelnen Rekonstruktion der Faktorenpaare $(\hat{C}_{:,i}, \hat{A}_{i,:})$ für $i = 1, \dots, s$. Kritisch anzumerken ist jedoch, dass das Zusammenspiel aller Komponenten der einzelnen Faktoren nicht berücksichtigt wird.

Exemplarisch für diesen Ansatz wird der in BTEM angewendete Algorithmus im Folgenden ausführlich erläutert und untersucht.

Auf dem Gebiet der *selbstmodellierenden Algorithmen* mit einer abgeschnittenen Singulärwertzerlegung als Ausgangspunkt und einem Anwendungsbezug zur Spektroskopie entsprechen die Ansätze und Algorithmen von Malinowski und hinter BTEM dem Stand der Forschung. Bei den klassischen Algorithmen zur Erzeugung von nichtnegativen Niedrigrangapproximationen ist der Algorithmus von Kim und Park eine geeignete Referenz.

5.3.2 Das Softwarepaket BTEM

Das kommerzielle Softwarepaket BTEM (*Band Target Entropy Minimization*) ist eine speziell zur Isolation einzelner Komponenten aus Spektroskopiedaten, gegeben in Form einer Matrix A , entwickelte Anwendung. Der zugehörige Algorithmus baut stark auf der von Malinowski in [63] genutzten Verfahrensweise auf, enthält jedoch einige sinnvolle beziehungsweise praktische Evolutionen und ist in einem Programmpaket für MATLAB inklusive graphischer Oberfläche umgesetzt. Als Verweise seien beispielhaft [11, 68, 86, 88] angegeben.

Die Problemklasse und der Lösungsansatz

Hinter der Software steht ein Algorithmus zum Erzeugen einer regularisierten nichtnegativen Niedrigrangapproximation einer Matrix A . Das Hauptaugenmerk der Niedrigrangapproximation $A \approx \hat{C}\hat{A}$ liegt auf dem Faktor \hat{A} , dessen Zeilen, ähnlich wie bereits von Malinowski entwickelt, einzeln bestimmt werden. Basis dieses Vorgehens ist die Annahme, dass sich die Einträge in bestimmten Indexbereichen von A hauptsächlich aus lediglich einem Teilfaktor $\hat{A}_{i_0,:}$ bestimmen. Diesen gilt es zu rekonstruieren. Dazu wird manuell ein zusammenhängender Bereich in den Indizes $I \subset \{1, \dots, n\}$ vorgegeben und eine „geeignete“ Linearkombination der rechtsseitigen Singulärvektoren von A bestimmt. Diese ergibt $\hat{A}_{i_0,:}$ und soll A auf I (demzufolge $A_{:,I}$) möglichst gut approximieren.

Der Programmaufbau

Die Anwendung von BTEM gliedert sich in die folgenden vier Schritte.

1. Zunächst werden die Messdaten eingelesen.
2. Anschließend folgt die Bestimmung einer Singulärwertzerlegung und die Betrachtung der rechtsseitigen Singulärvektoren. Im Zuge dessen werden bis zu 6 jeweils zusammenhängende Indexbereiche (*Band-Targets*, siehe Definition 5.6) festgelegt, mit dem Ziel diesbezüglich bis zu 6 Teilfaktoren $\hat{A}_{i,:}$, $i = 1, \dots, s \leq 6$, zu bestimmen.
3. In Stufe 3 erfolgt die konkrete Auswahl der Regularisierungsfunktionen f_i aus (2.7), speziell für den vorliegenden Datensatz. Die Minimierung der Zielfunktion aus (2.7) wird mittels eines *Simulated Annealing*-Algorithmus, [15], ausgeführt.

4. Der Abschluss umfasst die Auswertung der Ergebnisse, das heißt der Zeilen von \hat{A} . Die Gewichte \hat{C} werden nicht berechnet, lassen sich aber eigenständig mittels linearer Ausgleichrechnung bestimmen.

Approximation vom Rang 1

Ausgangspunkt ist zu einer Matrix $A \in \mathbb{R}^{k \times n}$, $\text{rank}(A) = s$, die idealisierte Faktorisierungsaufgabe der nichtnegativen Matrixfaktorisierung. Zu dem Faktor $\hat{A} \in \mathbb{R}^{s \times n}$, $\text{rank}(\hat{A}) = s$, ergibt sich mit

$$\hat{A}^+ = \hat{A}^T (\hat{A} \hat{A}^T)^{-1}$$

der zweite Faktor \hat{C} aus $A = \hat{C} \hat{A}$ als

$$\hat{C} = A \hat{A}^T (\hat{A} \hat{A}^T)^{-1}. \quad (5.13)$$

Diese Beziehung wird nun auf die Aufgabenstellung der nichtnegativen Niedrigrangapproximation und speziell der lokalen nichtnegativen Niedrigrangapproximation wie folgt übertragen.

Mit dem BTEM Algorithmus wird eine sukzessive Niedrigrangapproximation bestimmt. Das heißt das Faktorisierungsproblem wird in s Teilprobleme $A \approx \hat{C}_{:,i} \hat{A}_{i,:}$, $i = 1, \dots, s$, aufgespalten. Dabei wird die Auflösung aus (5.13) auf den Zusammenhang der Teilfaktoren eines Faktorenpaars $(\hat{C}_{:,1}, \hat{A}_{1,:})$ übertragen

$$\hat{C}_{:,1} = A \hat{A}_{1,:}^T (\hat{A}_{1,:} \hat{A}_{1,:}^T)^{-1} \quad (5.14)$$

und der Teilfaktor $\hat{A}_{1,:}$ als Linearkombination der ersten rechtsseitigen Singulärvektoren der Matrix A bestimmt

$$\hat{A}_{1,:} = T_{1,:} V_{:,1:z}^T \quad (5.15)$$

mit $s \leq z \leq \min\{k, n\}$.

Für das Residuum der Matrixfaktorisierung gilt nach diesem vereinfachenden Ansatz

$$\|A - \hat{C}_{:,1} \hat{A}_{1,:}\|_2^2 \geq \sigma_2^2, \quad \|A - \hat{C}_{:,1} \hat{A}_{1,:}\|_F^2 \geq \sum_{i=2}^{\min\{k,n\}} \sigma_i^2. \quad (5.16)$$

Der Ansatz einer lokalen Niedrigrangapproximation in BTEM

Die lokale Niedrigrangapproximation erfolgt in BTEM über Zielbereiche (*Band-targets*).

Definition 5.6. (*Zielbereiche, Band-targets*) Für die Variable x aus Faktorisierungsaufgabe 2.2 wird ein zusammenhängendes Intervall bestimmt, welches, übertragen auf das Frequenzraster $\bar{x} \in \mathbb{R}^n$, eine zusammenhängende Indexmenge I ergibt. Für diese Indexmenge wird eine lokale Niedrigrangapproximation bestimmt.

Abbildung 5.1 zeigt die Auswahl von fünf Bereichen für den Datensatz aus Anhang B.4.

Bemerkung 5.7. In den vorliegenden Veröffentlichungen ist nichts über die algorithmische Einarbeitung der Zielbereiche ausgesagt. Insofern kann eine Beurteilung der angeführten Rechenschritte nicht erfolgen und es lassen sich keine Aussagen bezüglich möglicher Probleme durch Überlagerungen und Coexistenzen der Teilfaktoren $\hat{A}_{i,:}$, $i = 1, \dots, s$, in den Zielbereichen treffen.

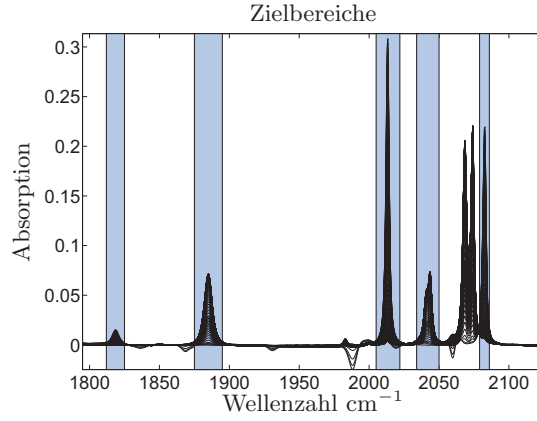


Abbildung 5.1: Auswahl von fünf Zielbereichen für den Datensatz aus Anhang B.4.

Bemerkung 5.8. Die einzelnen Teilfaktoren $(\hat{C}_{:,i}, \hat{A}_{i,:})$ werden jeweils aus den Originaldaten in A extrahiert. Das heißt $\hat{C}_{:,2}$ und $\hat{A}_{2,:}$ werden anhand von A und nicht $A - \hat{C}_{:,1}\hat{A}_{1,:}$ bestimmt.

Bemerkung 5.9. Der Ansatz die Faktorenpaare $(\hat{C}_{:,i}, \hat{A}_{i,:})$ einzeln abzuspalten ist sehr kritisch zu bewerten. Sei der idealisierte Fall einer Matrix A vom Rang s betrachtet. Der Ansatz aus (5.14) und (5.15) liefert nach Lemma 3.24 zwar eine Rangreduktion $\text{rank}(A - \hat{C}_{:,i}\hat{A}_{i,:}) = s - 1$, es wird jedoch ohne jede Grundlage unterstellt, dass für die verbleibende Matrix $A - \hat{C}_{:,i}\hat{A}_{i,:}$ exakte Zerlegungen in nichtnegative Faktoren vollen Ranges existieren und dass, falls es welche gibt, darunter zudem noch mindestens eine physikalisch-chemisch sinnvolle ist.

Die Optimierungsaufgabe

Die Nichtnegativitätsrestriktionen werden durch Strafterme eingebunden. Als weitere Regularisierungen werden in diskretisierter Form unter anderem das Betragsintegral des Faktors $\hat{a}_1(x)$ über x sowie die Norm der m -ten Ableitung ($m = 1, 2, 4$) von $\hat{a}_1(x)$ angewendet. Zusätzlich wird über (5.18) eine spezielle Form der in Abschnitt 5.1 kurz erläuterten Shannon-Entropie aus (5.3), vergleiche auch [76], bezüglich des Faktors $\hat{A}_{1,:}$ eingesetzt.

Nach [86] lautet das Minimierungsproblem insgesamt

$$\begin{aligned} F(\hat{C}, \hat{A}) &\rightarrow \min, & F &: \mathbb{R}^{k \times 1} \times \mathbb{R}^{1 \times n}, \\ F(\hat{C}, \hat{A}) &= \delta_{\text{Sh}} F_{\text{Sh}}(\hat{A}) + \delta_{\text{Der}} F_{\text{Der}}(\hat{A}) + \delta_{\text{Int}} F_{\text{Int}}(\hat{A}) + \gamma_1 P_{\hat{A}}(\hat{A}) + \gamma_2 P_{\hat{C}}(\hat{C}) \end{aligned} \quad (5.17)$$

mit den Regularisierungs- und Straffunktionen

$$\begin{aligned} F_{\text{Sh}}(\hat{A}) &= - \sum_i \frac{|(F_{\text{der},i}(\hat{A}))|}{F_{\text{Der}}(\hat{A})} \ln \left(\frac{|(F_{\text{der},i}(\hat{A}))|}{F_{\text{Der}}(\hat{A})} \right), \\ F_{\text{Der}}(\hat{A}) &= \sum_i \left| (F_{\text{der},i}(\hat{A})) \right|, \\ F_{\text{Int}}(\hat{A}) &= \sum_{i=1}^n |\hat{A}_i|, \\ P_{\hat{A}}(\hat{A}) &= F_1(\hat{A}) \begin{cases} 0 & F_1(\hat{A}) < \lambda_1, \\ 10 & \lambda_1 \leq F_1(\hat{A}) < \lambda_2, \\ 10^4 & \lambda_2 \leq F_1(\hat{A}) \end{cases} + \begin{cases} 0 & \max\{\hat{A}\} \leq \lambda_3, \\ 10^4 & \max\{\hat{A}\} > \lambda_3 \end{cases}, \\ P_{\hat{C}}(\hat{C}) &= 10^3 \sum_{i=1}^k \min(0, \hat{C})^2, \end{aligned} \quad (5.18)$$

den Hilfsfunktionen

$$F_{\text{der},i}(\hat{A}) \approx \frac{d^m \hat{A}}{dx^m}(\bar{x}_i), \quad m = 1, 2 \text{ oder } 4,$$

$$F_1(\hat{A}) = \sum_{i=1}^n \min(0, \hat{A}_i)^2,$$

den Parametern $\lambda_1 = 10^{-3}$, $\lambda_2 = 10^{-2}$ sowie dem variablen Parameter $\lambda_3 > 0$. Die Wahl von $\delta_{\text{Sh}}, \delta_{\text{Der}}, \delta_{\text{Int}} \in \{0, 1\}$ erfolgt manuell.

Bemerkung 5.10. Die Funktion $P_{\hat{A}}$ ist nur für Parametersätze mit $\lambda_1, \lambda_2 \in \{-\infty, 0, \infty\}$, $\lambda_1 \leq \lambda_2$, und $\lambda_3 \in \{-\infty, \infty\}$ stetig sowie dann auch differenzierbar.

Bemerkung 5.11. Der Einsatz der Straffunktion $P_{\hat{C}}$ ist unter Anwendung von (5.13) insofern kritisch zu bewerten, dass sie unter Umständen wenig Einfluss hat. Ist die Ausgangsmatrix A komponentenweise nichtnegativ und die Iterierte \hat{A} ebenfalls, so ist \hat{C} mit der Berechnung durch (5.13) auch komponentenweise nichtnegativ und es gilt $P_{\hat{C}}(\hat{C}) = 0$.

Eine Parameterreduktion durch die Singulärwertzerlegung

In der oben beschriebenen Form ist die Optimierungsfunktion über die Vektoren \hat{C} und \hat{A} eine Abbildung $\mathbb{R}^k \times \mathbb{R}^n \rightarrow \mathbb{R}$. Wird \hat{A} als Element des Bildraums der ersten z rechtsseitigen Singulärvektoren unter Verwendung von (5.15) bestimmt und \hat{C} in Abhängigkeit davon mittels (5.14), so vereinfacht sich das Optimierungsproblem zu

$$G(T) \rightarrow \min,$$

$$G : \mathbb{R}^{1 \times z} \rightarrow \mathbb{R}, \quad G(T) = F(\|T\|_2^{-2} U \Sigma_{:,1:z} T^T, T V_{:,1:z}),$$

wobei $U \Sigma V^T$ eine Singulärwertzerlegung von A ist.

5.3.3 Kritische Auseinandersetzung mit der lokalen Auflösung

Die Strategie des BTEM-Algorithmus ist es, nur eine lokale Auflösung der Matrixfaktorisierung vorzunehmen. Die Ungleichungen aus (5.16) zeigen, dass die Approximation durch ein dyadisches Produkt als Niedrigrangapproximation kritisch zu bewerten ist. Weiter soll an dieser Stelle kurz die Vorgehensweise aus (5.14) untersucht werden. Ebenso ist das Verhalten der Linearkombination aus (5.15) für den Fall interessant, dass eine Teilmatrix rekonstruiert werden soll, welche tatsächlich nur den Rang eins hat.

Kritische Auseinandersetzung des Zusammenhangs aus (5.14)

Der Ansatz $\hat{C}_{:,1}$ über die Linearkombination aus (5.14) als

$$\hat{C}_{:,1} \approx \tilde{C}_{:,1} = A \hat{A}_{1,:}^+$$

zu approximieren, ist lediglich in Sonderfällen mehr als nur eine grobe Näherung und liefert im Normalfall eine Überschätzung von $\hat{C}_{:,1}$. Eine Abschätzung für den Fehler liefert der folgende Satz.

Satz 5.12. Gegeben sei das Matrixprodukt $A = \hat{C} \hat{A}$ mit $A \in \mathbb{R}^{k \times n}$, $\hat{C} \in \mathbb{R}_+^{k \times s}$, $\hat{A} \in \mathbb{R}_+^{s \times n}$, $\text{rank}(A) = s$. Weiter sei zu $\hat{A}_{1,:}$ der Teilfaktor $\tilde{C}_{:,1}$ mittels (5.14) bestimmt als

$$\tilde{C}_{:,1} = A \hat{A}_{1,:}^+.$$

Dann gilt

$$\tilde{C}_{:,1} = \hat{C}_{:,1} \Leftrightarrow \hat{A}_{1,:} \perp \hat{A}_{i,:}, \quad \forall i = 2, \dots, s, \quad (5.19)$$

$$\tilde{C}_{i,1} \geq \hat{C}_{i,1}, \quad i = 1, \dots, k, \quad (5.20)$$

$$\max_{i=2,\dots,s} \hat{A}_{i,:}^T \hat{A}_{1,:} \|\hat{C}_{:,i}\| \leq \|\hat{A}_{1,:}\|_2^2 \|\tilde{C}_{:,1} - \hat{C}_{:,1}\| \leq \sum_{i=2}^s \hat{A}_{i,:}^T \hat{A}_{1,:} \|\hat{C}_{:,i}\|. \quad (5.21)$$

Beweis: Aus $\text{rank}(A) = s$ folgt $\|\hat{A}_{1,:}\|_2 > 0$. Weiter gilt

$$\begin{aligned} \tilde{C}_{:,1} &= \left(\sum_{i=1}^s \hat{C}_{:,i} \hat{A}_{i,:} \right) \hat{A}_{1,:}^+, \\ \tilde{C}_{:,1} - \hat{C}_{:,1} &= \left(\sum_{i=2}^s \hat{C}_{:,i} \hat{A}_{i,:} \right) \hat{A}_{1,:}^T \|A_{1,:}\|_2^{-2} \end{aligned}$$

und es folgt die rechte Ungleichung von (5.21). Aus $\text{rank}(\hat{C}) = s$ folgt (5.19). Weiter sind alle Einträge in \hat{C} und \hat{A} nichtnegativ und es folgt (5.20) sowie die linke Ungleichung von (5.21). \square

Speziell die untere Schranke der Abschätzung für $\|\tilde{C}_{:,1} - \hat{C}_{:,1}\|$ aus (5.21) beinhaltet starke Voraussetzungen an den Faktor \hat{A} , um eine gute Näherung für $\tilde{C}_{:,1}$ zu erhalten.

In Abschnitt 6.5 wird eine kleine Modifikation der Vorgehensweise aus (5.14) vorgestellt, welche die Fehlerschranken aus (5.21) deutlich herabsetzt. Dabei beschränken sich die Skalarprodukte $(\hat{A}_{i,:}, \hat{A}_{1,:})$ der Fehlerabschätzungen aus (5.21) lokal auf die Indexmenge I der Zielbereiche

$$\max_{i=2,\dots,s} \hat{A}_{i,I}^T \hat{A}_{1,I} \|\hat{C}_{:,i}\| \leq \|\hat{A}_{1,I}\|_2^2 \|\tilde{C}_{:,1} - \hat{C}_{:,1}\| \leq \sum_{i=2}^s \hat{A}_{i,I}^T \hat{A}_{1,I} \|\hat{C}_{:,i}\|,$$

vergleiche dazu Satz 6.25.

Ergebnis der lokalen Niedrigrangapproximation einer Teilmatrix mit Rang 1

Satz 5.12 zeigt, wie sich der Fehler der Approximation für $\hat{C}_{:,1}$ aus (5.14) verhält. Ein weiterer Kritikpunkt der lokalen Rekonstruktion setzt bei dem Fall an, dass eine Teilmatrix $A_{:,I}$ aus Spalten von A durch $\hat{A}_{1,I}$ approximiert werden soll, wobei sich $A_{:,I}$ tatsächlich als dyadisches Produkt $\tilde{C}_{:,1} \hat{A}_{1,I}$ darstellen lässt.

Sei dazu für die idealisierte Faktorisierungsaufgabe 2.3 der Fall angenommen, es gibt eine nicht-leere zusammenhängende Menge $I \subset \{1, \dots, n\}$, so dass $\text{rank}(A_{:,I}) = 1$ gilt. Dann folgt für den idealisierten Fall mit der Beziehung $\hat{A} = TV_{:,1:s}^T$ aus (5.15), dass ebenfalls $\text{rank}(V_{I,1:s}) = 1$ sein muss. Insgesamt ergibt sich so zur Bestimmung von $\hat{A}_{1,:}$ auf der Indexmenge I ein eindimensionaler Vektorraum für $\hat{A}_{1,I}$.

Fazit 5.13. Für den oben beschriebenen Fall konzentriert sich die Bestimmung des Teilfaktors $\hat{A}_{1,:}$ und damit auch dessen Regularisierung grundsätzlich auf den Indexbereich $\{1, \dots, n\} \setminus I$, also den Indexbereich, auf welchem A nicht lokal approximiert werden soll.

5.4 Kritische Zusammenfassung der vorgestellten Lösungsansätze

Die in diesem Kapitel vorgestellten Methoden eignen sich aus verschiedenen Gründen nur bedingt zur Lösung der Faktorisierungsaufgaben 2.3, 2.4 (nichtnegative Matrixfaktorisierung, Niedrigrangapproximation) und 2.5 (regularisierte nichtnegative Niedrigrangapproximation) mit Bezug zu spektroskopischen Daten. Dazu seien die Methoden abschließend kurz zusammengefasst.

Hauptkomponentenanalyse

Das grundsätzliche Problem der Hauptkomponentenanalyse sind die nicht beachteten Nichtnegativitätsrestriktionen an die Faktoren. Zudem konzentriert sich die Faktorisierung nur auf einen der Faktoren X oder Y und die Bestimmung untereinander unkorrelierter Teilfaktoren. Der zweite Faktor bleibt hingegen unberücksichtigt, was der zu lösenden Faktorisierungsaufgabe grundsätzlich widerspricht. Der Ansatz einer Faktorisierung in untereinander unkorrelierte Teilfaktoren für einen der Faktoren kann als durchaus sinnvoll angesehen werden, wobei dafür die Problemklasse der Daten entscheidend ist.

Durch zusätzliche Nichtnegativitätsrestriktionen ließe sich diese Methode auch für die vorliegenden Faktorisierungsaufgaben anwenden.

Unabhängigkeitsanalyse

Die Kritikpunkte aus der Hauptkomponentenanalyse lassen sich direkt auf die Unabhängigkeitsanalyse übertragen. Der Ansatz untereinander möglichst unabhängige Teilfaktoren für einen der Faktoren zu erzeugen, ist im Vergleich zur Hauptkomponentenanalyse jedoch für einige Anwendung ein überzeugenderer Zugang zur Bestimmung einer Faktorisierung.

Nichtnegative Matrixfaktorisierung

Die vorgestellten klassischen Algorithmen zur nichtnegativen Matrixfaktorisierung entsprechen dem Anforderungsprofil zur Lösung der vorliegenden Faktorisierungsaufgaben. Jedoch bietet einerseits die in Abschnitt 5.2 erwähnte Variante von Lee und Seung überhaupt keine zusätzlichen Regularisierungsmöglichkeiten sowie andererseits die vorgestellte Methode von Kim und Park nicht genug Flexibilität hinsichtlich der Regularisierungen, welche lediglich über α und β steuerbar sind. Weiter wäre es bei der Anwendung auf Daten mit vergleichsweise kleiner Störung E sinnvoll, die Faktoren X und Y aus speziell gewählten Unterräumen des $\mathbb{R}^{k \times s}$ und des $\mathbb{R}^{s \times n}$ niedriger Dimension zu bestimmen. Dies würde den Optimierungsaufwand reduzieren, vergleichbar mit dem in Abschnitt 4.3 beschriebenen Ansatz. Da es sich bei den Verfahren jedoch um einen allgemein gehaltenen Ansatz zur Lösung von Aufgaben der Niedrigrangapproximation handelt, beschränkt sich dieser Kritikpunkt auf die in dieser Arbeit betrachtete spezielle Problemklasse.

Der Algorithmus BTEM

Der Algorithmus BTEM beinhaltet eine deutliche Reduktion des Aufwandes. Die Faktoren werden nicht autonom bestimmt, sondern über eine Transformation der Faktoren einer abgeschnittenen Singulärwertzerlegung von A . Somit werden die Spalten von X und Zeilen von Y als Elemente von niedrigdimensionalen Untervektorräumen, aufgespannt von den ersten z Singulärvektoren, berechnet. Die eingesetzten Regularisierungen sind, abgesehen von der in Bemerkung 5.10 benannten Unstetigkeit von $P_{\hat{A}}(\hat{A})$, für den ausgewählten Anwendungsbereich der optischen Spektroskopie geeignet. Sie sind diesem angepasst und flexibel.

Hauptkritikpunkte sind die lokale Niedrigrangapproximation mit einem dyadischen Produkt, welche der Problemstellung auch lokal gekoppelter Daten widerspricht (vergleiche (5.16) sowie Bemerkung 5.9), die Berechnung einer Approximation für $\hat{C}_{:,1}$ in Abhängigkeit von $\hat{A}_{1,:}$ (Abschätzung aus Satz 5.12) sowie die Ausweitung der lokal betrachteten Komponenten auf den gesamten Indexbereich (Fazit 5.13). Weiter werden die einzelnen Teilfaktoren in \hat{C} nicht von der Routine ausgegeben und müssen im Anschluss eigenständig mittels linearer Ausgleichsrechnung bestimmt werden.

Trotz der Kritik an den mathematischen Voraussetzungen bleibt festzuhalten, dass der BTEM-Algorithmus in der Praxis bei der Anwendung auf Daten der FT-IR Spektroskopie gute Ergebnisse liefert. Gegenüber Störungen verhält sich der Algorithmus robust.

6 Regularisierte nichtnegative Niedrigrangapproximationen und ihre Implementierung in der Methode Pure Component Decomposition

In Kapitel 3 ist gezeigt, dass der kontinuierliche Produktansatz unter bestimmten Voraussetzungen, von der Skalierung und Sortierung abgesehen, eindeutig lösbar ist. Dabei sind sowohl die notwendigen Anforderungen an die Ausgangsfaktoren aus Satz 3.6 als auch die hinreichenden Bedingungen aus Satz 3.8 sehr einschränkend und nur in Sonderfällen erfüllt.

Analoges gilt ebenfalls für die nichtnegative Matrixfaktorisierung; im Normalfall sind die Faktoren nicht eindeutig. Basieren die Daten A auf einer zu analysierende Modellierung, so besteht der Bedarf, aus allen möglichen Faktorisierungen bestimmte zu selektieren, um so tiefgreifende Aufschlüsse durch die Zerlegung zu erhalten. Eine Möglichkeit der Selektion sind Regularisierungen. Dies führt auf die Aufgabe der Berechnung einer regularisierten nichtnegativen Niedrigrangapproximation.

Übersicht

Dieses Kapitel bildet den praktischen Kernteil der Arbeit; es wird ein neu entwickelter Algorithmus zur Lösung der Faktorisierungsaufgabe 2.5 vorgestellt und diskutiert.

In den Abschnitten 6.1 und 6.2 werden die in dem Algorithmus eingesetzten Regularisierungen erläutert. Dabei sind in 6.1 allgemeine Regularisierungen behandelt und in 6.2 eine sehr spezielle; es ist einer der Faktoren mit einem Anfangswertproblem verknüpft. Die sich ergebende Zielfunktion liegt in Form eines nichtlinearen Quadratmittelproblems vor, dessen Lösung kurz beschrieben wird. Ausgangspunkt für die Niedrigrangapproximation ist der kontinuierliche Produktansatz aus (2.2) mit stetigen Funktionen $w_i(x)$, $v_i(y)$, $i = 1, \dots, s$. Für den diskretisierten Fall bietet dies die Möglichkeit einer stufenweisen Ausdünnung der Daten in unterschiedlichen Grobgridern. Diese Vorgehensweise einer Multi-Level-Strategie ist in 6.4 erläutert. Weiter wird in 6.5 ein nicht zu empfehlender Algorithmus einer lokalen Niedrigrangapproximation eingeführt und analysiert. Eine Analyse der sensiblen Wahl der Gewichtungparameter der Regularisierungsfunktionen erfolgt in Abschnitt 6.6.

Insgesamt entstand aus diesen Ideen ein Programmpaket, welches auf die Niedrigrangapproximation von spektroskopischen Daten ausgelegt ist. Die Methode trägt den Namen *Pure Component Decomposition*, kurz PCD.

Notation und Aufgabenstellung

Ebenso wie in Abschnitt 5.3 werden, der üblichen Notation der Chemie folgend, in diesem Kapitel die Bezeichnungen $\hat{C} = X$, $\hat{A} = Y$ verwendet. Für den kontinuierlichen Produktansatz 2.2 werden die Variablen mit $t = x$, $x = y$ und die Funktionen mit $\hat{c} = w$, $\hat{a} = v$ bezeichnet.

Die Aufgabenstellung lautet demnach wie folgt.

Faktorisierungsaufgabe 6.1. *Zu einer Matrix $A \in \mathbb{R}^{k \times n}$, $A_{i,j} \geq -\varepsilon$ mit $0 \leq \varepsilon \ll \max_{i,j} A_{i,j}$ und einer Zahl $s \in \mathbb{N}$, $s \leq \text{rank}(A)$, ist eine regularisierte nichtnegative Niedrigrangapproximation*

$$\frac{1}{2} \|A - \hat{C}\hat{A}\|_F^2 + \frac{1}{2} \sum_{i=1}^p \gamma_i^2 f_i(\hat{C}, \hat{A}) \rightarrow \min, \quad (6.1)$$

durch $\hat{C} \in \mathbb{R}_+^{k \times s}$ und $\hat{A} \in \mathbb{R}_+^{s \times n}$, $\text{rank}(\hat{C}\hat{A}) = s$, gesucht mit den Regularisierungsfunktionen $f_i(\hat{C}, \hat{A}) \geq 0$, $i = 1, \dots, p$, sowie deren Gewichtung $\gamma_i \in \mathbb{R}$, $i = 1, \dots, p$.

Der angewendete Lösungsansatz

Ausgangspunkt sei zunächst eine Matrix $A \in \mathbb{R}_+^{k \times n}$, welche exakt in zwei nichtnegative Faktoren $\hat{C} \in \mathbb{R}_+^{k \times s}$ und $\hat{A} \in \mathbb{R}_+^{s \times n}$ jeweils vollen Ranges zerfällt. Wie in Abschnitt 4.3 erläutert, ist es möglich mit Hilfe einer regulären Matrix $T \in \mathbb{R}^{s \times s}$ und den ersten s links- und rechtsseitigen Singulärvektoren auf die Faktoren \hat{C} und \hat{A} zu schließen (inklusive der Skalierung durch die Singulärwerte)

$$\hat{C} = U\Sigma_{:,1:s}T^{-1}, \quad \hat{A} = TV_{:,1:s}^T.$$

Ein solcher Ansatz wird nun verwendet und es ist die Aufgabe eine geeignete Transformation T zu bestimmen. Bei der Berechnung einer Niedrigrangapproximation gestörter Daten wird eine Matrix $T \in \mathbb{R}^{s \times z}$ angewendet

$$\hat{C} = U\Sigma_{:,1:z}T^+, \quad \hat{A} = TV_{:,1:z}^T. \quad (6.2)$$

Vorüberlegung zur Skalierung der Faktoren

Einige der später vorgestellten Regularisierungsfunktionen besitzen Eigenschaften wie etwa

$$f(\alpha^{-1}\hat{C}, \alpha\hat{A}) = \alpha^2 f(\hat{C}, \hat{A}),$$

so dass der unbekannte Skalierungsfaktor der einzelnen Faktorenpaare bei der Optimierung hinderlich ist. Eine Möglichkeit diesem Problem entgegenzutreten, ist eine, die Allgemeinheit nicht einschränkende, Normierung der Zeilen von \hat{A}

$$\hat{C}_{:,i} \leftarrow \|\hat{A}_{i,:}\|_\infty \hat{C}_{:,i}, \quad i = 1, \dots, s, \quad (6.3)$$

$$\hat{A}_{i,:} \leftarrow \|\hat{A}_{i,:}\|_\infty^{-1} \hat{A}_{i,:}, \quad i = 1, \dots, s, \quad (6.4)$$

auf 1 vorzunehmen (also indirekt $T_{i,:} \leftarrow \|\hat{A}_{i,:}\|_\infty^{-1} T_{i,:}$) und so die Rechnung im Sinne der Skalierung auf Repräsentanten der Äquivalenzklassen zu beschränken.

Diese direkte Anwendung birgt jedoch ein Problem. Sei $T^{(\ell)} \in \mathbb{R}^{s \times z}$ die ℓ -te Iterierte bei der späteren Minimierung von F aus (6.11), wobei $T^{(\ell)}$ den vollen Rang habe. Wird eine Zeile von $T^{(\ell)}$ in eine Nullmatrix eingebettet, so ergibt sich für eine geeignete Änderung der Iterierten um diese Matrix eine im Rahmen der Maschinengenauigkeit gleiche Folgerechnung

$$F(T^{(\ell)} + \gamma \text{diag}(e_m)T^{(\ell)}) = F(T^{(\ell)}), \quad m = 1, \dots, s, \quad \gamma > -1.$$

Dabei ist $e_m \in \mathbb{R}^s$ der m -te Einheitsvektor.

Es existiert also ein s -dimensionaler Vektorraum V , so dass eine Variation von $T^{(\ell)}$ um geeignet skalierte Elemente aus V auf eine gleiche Folgerechnung wie $T^{(\ell)}$ selbst führt. Für eine mit $\phi : \mathbb{R}^{sz} \rightarrow \mathbb{R}^{s \times z}$,

$$(\phi(x))_{i,j} = x_{(i-1)z+j}, \quad i = 1, \dots, s, \quad j = 1, \dots, z, \quad (6.5)$$

aus $F : \mathbb{R}^{s \times z} \rightarrow \mathbb{R}$ erzeugte Funktion $G : \mathbb{R}^{sz} \rightarrow \mathbb{R}$,

$$G(x) = F(\phi(x))$$

besitzt die Hessematrix $H_G(x)$ aller zweiten partiellen Ableitungen damit für alle $x \in \mathbb{R}^{sz}$ mit $\text{rank}(\phi(x)) = s$ stets einen s -dimensionalen Kern. Es kann sich somit das hinreichende Kriterium einer strikten lokalen Minimalstelle x^* (positiv definite Hessematrix) nicht einstellen.

Eine praktisch bessere Umsetzung ist der Weg über eine Straffunktion

$$f_1 : \mathbb{R}^{s \times n} \rightarrow \mathbb{R}, \quad f_1(\hat{A}) = \sum_{i=1}^s (1 - \|\hat{A}_{i,:}\|_{\infty})^2. \quad (6.6)$$

6.1 Die Straf- und Regularisierungsfunktionen

In diesem Teilabschnitt werden die in dem Faktorisierungsalgorithmus eingesetzten Regularisierungsfunktionen f_i zur Lösung der Faktorisierungsaufgabe 2.5 vorgestellt. Zusätzlich zur Approximationsforderung werden zunächst die Nichtnegativitätsrestriktionen der Faktoren \hat{C}, \hat{A} in Straffunktionen eingearbeitet. Neben diesen Problemrestriktionen werden geeignete Regularisierungsfunktionen motiviert und eingeführt. Insgesamt werden alle Funktionen zunächst für das Optimierungsproblem aus (6.1), also für \hat{C}, \hat{A} direkt, eingeführt. Die Transformationen aus (6.2) werden erst zum Schluss angewendet.

Niedrigrangapproximation und lokale Niedrigrangapproximation

Die Approximationsforderung aus (6.1) wird aufgrund des einfacheren Umgangs mit allen Straf- und Regularisierungsfunktionen als

$$f_0 : \mathbb{R}^{k \times s} \times \mathbb{R}^{s \times n} \rightarrow \mathbb{R}, \quad f_0(\hat{C}, \hat{A}) = \|A - \hat{C}\hat{A}\|_F^2$$

eingeführt. Für eine lokale Niedrigrangapproximation von A nur auf der Zeilenindexmenge $I \subset \{1, \dots, n\}$ ergibt sich

$$f_0(\hat{C}, \hat{A}) = \|A_{:,I} - \hat{C}\hat{A}_{:,I}\|_F^2.$$

Die Nichtnegativitätsrestriktionen

Für die idealisierte nichtnegative Matrixfaktorisierung lauten die Forderungen

$$\begin{aligned} \hat{C}_{i,j} &\geq 0, & i = 1, \dots, k, & \quad j = 1, \dots, s, \\ \hat{A}_{i,j} &\geq 0, & i = 1, \dots, s, & \quad j = 1, \dots, n, \end{aligned}$$

an die Faktoren \hat{C}, \hat{A} . Für gestörte Daten und die Aufgabe der Niedrigrangapproximation ist es mitunter sinnvoll, Werte $a_{\min}, c_{\min} \leq 0$ als untere Schranken einzuführen. An dieser Stelle wirkt die Restriktion der Normierung von \hat{A} aus (6.6) einer ungewollten Ausnutzung einer möglichen freien Skalierung bei der Anwendung absoluter Schranken entgegen.

Die Nichtnegativitätsrestriktionen werden in zwei Straffunktionen eingebettet

$$\begin{aligned} f_2 : \mathbb{R}^{s \times n} &\rightarrow \mathbb{R}, & f_2(\hat{A}) &= \sum_{i=1}^s \sum_{j=1}^n (\varphi(\hat{A}_{i,j}))^2, \\ f_3 : \mathbb{R}^{k \times s} &\rightarrow \mathbb{R}, & f_3(\hat{C}) &= \sum_{i=1}^k \sum_{j=1}^s (\psi(\hat{C}_{i,j}))^2 \end{aligned}$$

mit $\varphi, \psi : \mathbb{R} \rightarrow \mathbb{R}$ definiert als

$$\varphi(a) = \begin{cases} a - a_{\min} & \text{für } a_{\min} > a, \\ 0 & \text{für } a_{\min} \leq a \end{cases}, \quad \psi(c) = \begin{cases} c - c_{\min} & \text{für } c_{\min} > c, \\ 0 & \text{für } c_{\min} \leq c. \end{cases}$$

Bemerkung 6.2. Die Konstante c_{\min} ist stark von den Daten in A abhängig. Unter Umständen ist es sinnvoll für f_3 eine Normierung der Spalten von \hat{C} bezüglich der Maximumsnorm vorzunehmen.

Eine problemabhängige Modellkorrektur

Unter besonderen Umständen ist eine leichte Modellkorrektur in dem Sinne nötig, dass die Nichtnegativitätsrestriktionen für einen Teilfaktor $C_{:,j_0}$ aufgehoben werden. Dies führt auf die Faktorisierungsaufgabe 6.3.

Faktorisierungsaufgabe 6.3. Zu einer gegebenen Matrix $A \in \mathbb{R}^{k \times n}$, einer Zahl $s \in \mathbb{N}$, $s \leq \text{rank}(A)$, und einem Index $j_0 \in \{1, \dots, s\}$ ist eine regularisierte Niedrigrangapproximation

$$\frac{1}{2} \|A - \hat{C} \hat{A}\|_F^2 + \frac{1}{2} \sum_{i=1}^p \gamma_i^2 f_i(\hat{C}, \hat{A}) \rightarrow \min,$$

$$\hat{A} \in \mathbb{R}_+^{s \times n}, \quad \hat{C} \in \mathbb{R}^{k \times s}, \quad \hat{C}_{i,j} \geq 0, \quad i = 1, \dots, k, \quad j = 1, \dots, s, \quad j \neq j_0,$$

gesucht.

Eine derartige Faktorisierungsaufgabe führt auf eine Modifikation der Straffunktion $f_3(\hat{C})$ zu

$$\tilde{f}_3 : \mathbb{R}^{k \times s} \rightarrow \mathbb{R}, \quad \tilde{f}_3(\hat{C}) = \sum_{i=1}^k \sum_{j=1}^s \left((1 - \delta_{j,j_0}) \tilde{\psi}_1(\hat{C}_{i,j}) + \delta_{j,j_0} \tilde{\psi}_2(\hat{C}_{i,j}) \right)^2 \quad (6.7)$$

mit $\tilde{\psi}_1, \tilde{\psi}_2 : \mathbb{R} \rightarrow \mathbb{R}$ definiert als

$$\begin{aligned} \tilde{\psi}_1(c) &= \begin{cases} c - c_{\min} & \text{für } c_{\min} > c, \\ 0 & \text{für } c_{\min} \leq c \leq c_{\max}, \\ c - c_{\max} & \text{für } c > c_{\max} \end{cases}, \\ \tilde{\psi}_2(c) &= \begin{cases} c + c_{\max} & \text{für } -c_{\max} > c, \\ 0 & \text{für } -c_{\max} \leq c \leq c_{\max}, \\ c - c_{\max} & \text{für } c > c_{\max}. \end{cases} \end{aligned}$$

Für die idealisierte Faktorisierungsaufgabe 2.3 ergibt sich mit der Normierung $\|\hat{A}_{i,:}\|_{\infty} = 1$, $i = 1, \dots, s$, und aus der Aufgabenstellung $A_{i,j} \geq 0$ die Beschränkung

$$\hat{C}_{i,j} \leq \max_{i,j} A_{i,j}, \quad i = 1, \dots, k, \quad j = 1, \dots, s,$$

welche sich mit dem Einfluss von leichten Störungen auch auf die Faktorisierungsaufgabe 2.5 überträgt. Für die Modellkorrektur aus Faktorisierungsaufgabe 6.3 folgt dies nicht und um etwaigen unerwünschten Egalisierungseffekten Einhalt zu gebieten, wird dies direkt in der Restriktion (6.7) über c_{\max} für den Faktor \hat{C} gefordert (vergleiche dazu beispielsweise in Abbildung B.10 die Faktorisierung durch den Algorithmus BTEM).

6.1.1 Regularisierungen der Faktoren

Wie bereits zum Anfang dieses Abschnittes angemerkt, sind die in Kapitel 3 angeführten notwendigen und hinreichenden Kriterien für die Existenz nur einer Äquivalenzklasse an Faktorisierungen sehr einschränkend. Außer in Spezialfällen lassen allein die Nichtnegativitätsbedingungen und die Faktorisierungsforderung beliebig viele Äquivalenzklassen an Auflösungen und Niedrigrangapproximationen zu.

Ziel der Regularisierungen aus Faktorisierungsaufgabe 2.5 ist es, unter all diesen Faktorenpaaren zu selektieren. Die Wahl dieser Regularisierungen ist dabei jeweils spezifisch für das der Matrix zugrunde liegende Anwendungsfeld.

Glatte und dünn besetzte Faktoren

Eine einfache Form der Regularisierung sind Glattheitsforderungen für die einzelnen Faktoren, beispielsweise über die erste und die zweite Ableitung. Ebenso lässt sich eine Regularisierung mittels des Betragsintegrals anwenden. Für den kontinuierlichen Produktansatz aus Faktorisierungsaufgabe 2.2 führen diese Ansätze auf die Regularisierungsfunktionen

$$\begin{aligned} \tilde{f}_4(\hat{a}) &= \sum_{i=1}^s \int \left(\frac{d\hat{a}_i}{dx}(x) \right)^2 dx, & \tilde{f}_6(\hat{a}) &= \sum_{i=1}^s \int \hat{a}_i^2(x) dx, \\ \tilde{f}_5(\hat{a}) &= \sum_{i=1}^s \int \left(\frac{d^2\hat{a}_i}{dx^2}(x) \right)^2 dx, & \tilde{f}_7(\hat{c}) &= \sum_{i=1}^s \int \left(\frac{d\hat{c}_i}{dt}(t) \right)^2 dt. \end{aligned}$$

Die Übertragung auf das diskrete Problem führt auf die lineare Tichonov-Regularisierung, [5, 19, 36, 61], welche sich für einen Vektor $x \in \mathbb{R}^n$ als

$$T_\alpha(x) = \alpha \|Bx\|^2$$

mit einer linearen Abbildung $B \in \mathbb{R}^{m \times n}$, $m \leq n$, und $\alpha > 0$ ergibt. Sei dabei das Diskretisierungsgitter für beide Variablen t und x jeweils äquidistant mit den Schrittweiten τ beziehungsweise h . Die angewendeten Regularisierungsfunktionen lauten

$$f_4, f_5 : \mathbb{R}^{s \times n} \rightarrow \mathbb{R}, \quad f_i(\hat{A}) = \sum_{j=1}^s \left\| B^{(i)} \hat{A}_{j,:}^T \right\|_{\text{diag}(b^{(4)})}^2, \quad i = 4, 5, \quad (6.8)$$

$$f_6 : \mathbb{R}^{s \times n} \rightarrow \mathbb{R}, \quad f_6(\hat{A}) = \sum_{i=1}^s \left\| \hat{A}_{i,:}^T \right\|_{\text{diag}(b^{(6)})}^2, \quad (6.9)$$

$$f_7 : \mathbb{R}^{z \times k} \rightarrow \mathbb{R}, \quad f_7(\hat{C}) = \sum_{i=1}^s \left\| B^{(7)} \hat{C}_{:,i} \right\|_{\text{diag}(b^{(7)})}^2$$

mit den Matrizen

$$B^{(4)} \in \mathbb{R}^{(n-2) \times n}, \quad (B^{(4)})_{ij} = \frac{1}{2h} \begin{cases} -1 & \text{für } j - i = 0, \\ 1 & \text{für } j - i = 2, \\ 0 & \text{sonst} \end{cases},$$

$$B^{(5)} \in \mathbb{R}^{(n-2) \times n}, \quad (B^{(5)})_{ij} = \frac{1}{h^2} \begin{cases} 1 & \text{für } j - i = 0, \\ -2 & \text{für } j - i = 1, \\ 1 & \text{für } j - i = 2, \\ 0 & \text{sonst} \end{cases},$$

$$B^{(7)} \in \mathbb{R}^{(k-2) \times k}, \quad (B^{(7)})_{ij} = \frac{1}{2\tau} \begin{cases} -1 & \text{für } j - i = 0, \\ 1 & \text{für } j - i = 2, \\ 0 & \text{sonst} \end{cases},$$

den Vektoren

$$\begin{aligned} b^{(4)} &\in \mathbb{R}^{n-2}, & b^{(4)} &= h(0.5, 1, 1, \dots, 1, 0.5)^T \in \mathbb{R}^{n-2}, \\ b^{(6)} &\in \mathbb{R}^n, & b^{(6)} &= h(0.5, 1, 1, \dots, 1, 0.5)^T \in \mathbb{R}^n, \\ b^{(7)} &\in \mathbb{R}^{k-2}, & b^{(7)} &= \tau(0.5, 1, 1, \dots, 1, 0.5)^T \in \mathbb{R}^{k-2} \end{aligned}$$

und der durch eine symmetrisch positiv definite Matrix B induzierten Norm

$$\|x\|_B^2 = (x, Bx) = x^T Bx.$$

Lokal dicht besetzte Faktoren

Eine Möglichkeit der Regularisierung bei einer lokalen Niedrigrangapproximation ist es, die Integralberechnung von \tilde{f}_6 auf den Bereich außerhalb des betrachteten Approximationsbereichs I für x zu beschränken

$$\tilde{f}_8 = \tilde{f}_6(\hat{a}) - \sum_{i=1}^s \int_I \hat{a}_i^2(x) \, dx.$$

Die eingesetzte diskrete Variante lautet

$$f_8 : \mathbb{R}^{s \times z} \rightarrow \mathbb{R}, \quad f_8(T) = \sum_{i=1}^s \left\| \hat{A}_{i,:}^T \right\|_{\text{diag}(b^{(8)})}^2$$

mit einem geeigneten Vektor $b^{(8)}$ analog zu $b^{(6)}$, nur eingeschränkt auf den Bereich außerhalb des Intervalls I .

Systemrestriktionen

Durch Vorkenntnisse über Teile der Faktoren lässt sich die Menge aller Äquivalenzklassen möglicher Faktorisierungen auf eine Menge relevanter einschränken, siehe Abschnitt 4.5. Eine Möglichkeit a-priori Kenntnisse über \hat{C} und \hat{A} auf die Faktorisierung zu übertragen sind Regularisierungsfunktionen.

Da es sich einerseits sowohl bei den Einträgen in A , wie auch oftmals bei den bekannten Teilfaktoren um gestörte Daten handelt und andererseits die Skalierungsfrage bleibt, ist eine Anpassung erforderlich. Seien $c_0 \in \mathbb{N}$ Teilfaktoren von \hat{C} unskaliert bekannt und in $\bar{C} \in \mathbb{R}^{k \times c_0}$ vorgegeben. Die anzupassenden Spalten in \hat{C} seien beispielsweise $\hat{C}_{:,1:c_0}$. Dann liefert

$$f_9 : \mathbb{R}^{k \times c_0} \rightarrow \mathbb{R}, \quad f_9(\hat{C}) = \|\bar{C} - \hat{C}_{:,1:c_0} \Theta\|_F^2$$

mit der Diagonalmatrix

$$\Theta \in \mathbb{R}^{c_0 \times c_0}, \quad \Theta = \text{diag} \left((\hat{C}_{:,1}^T \bar{C}_{:,1}) / (\hat{C}_{:,1}^T \hat{C}_{:,1}), \dots, (\hat{C}_{:,c_0}^T \bar{C}_{:,c_0}) / (\hat{C}_{:,c_0}^T \hat{C}_{:,c_0}) \right)$$

ein Maß für die Abweichung der variablen Größe \hat{C} von der Vorgabe \bar{C} . Analoges lässt sich für \hat{A} durch

$$\begin{aligned} f_{10} : \mathbb{R}^{a_0 \times n} &\rightarrow \mathbb{R}, & f_{10}(\hat{A}) &= \|\bar{A} - \Theta \hat{A}_{1:a_0,:}\|_F^2, \\ \Theta &\in \mathbb{R}^{a_0 \times a_0}, & \Theta &= \text{diag} \left((\hat{A}_{1,:}^T \bar{A}_{1,:}) / (\hat{A}_{1,:}^T \hat{A}_{1,:}), \dots, (\hat{A}_{a_0,:}^T \bar{A}_{a_0,:}) / (\hat{A}_{a_0,:}^T \hat{A}_{a_0,:}) \right) \end{aligned}$$

definieren, wobei a_0 die Zahl der bekannten Teilfaktoren $\bar{A} \in \mathbb{R}^{a_0 \times n}$ ist.

Bemerkung 6.4. Bei einer simultanen Anwendung von f_9 und f_{10} ist zu beachten, dass die einzelnen Teilfaktoren eine Zuordnung erhalten müssen. In f_9 und f_{10} sind die bekannten Teilfaktoren von 1 bis c_0 beziehungsweise von 1 bis a_0 angeordnet. Da die beiden Anordnungen in der Regel nicht zueinander passen beziehungsweise die Mengen der betrachteten Teilfaktoren sogar disjunkt sind, ist eine andere Vorgabe der Ordnung nötig.

Weiter lassen sich Kenntnisse über die Anzahl der Teilfaktoren aus \hat{A} in bestimmten Linearkombinationen aus A einarbeiten. Sei dazu $I \subset \{1, \dots, k\}$ und $c^{(s)} \in \mathbb{N}^{|I|}$ gegeben, so dass die Forderung

$$\forall i \in I \exists J^{(i)} \subset \{1, \dots, s\}, |J^{(i)}| \geq c_i^{(s)} : \hat{C}_{i,j} = 0, \quad \forall j \in J^{(i)},$$

besteht. Als Regularisierungsfunktion zur Umsetzung dessen ist

$$f_{11} : \mathbb{R}^{k \times s} \rightarrow \mathbb{R}, \quad f_{11}(\hat{C}) = \sum_{i \in I} \sum_{j=1}^{c_i^{(s)}} \frac{\hat{C}_{i,\pi_j^{(i)}}^2}{\left\| \hat{C}_{:, \pi_j^{(i)}} \right\|_\infty^2} \quad (6.10)$$

gewählt. Dabei werden die Koeffizienten aus $\hat{C}_{i,:}$ relativ betrachtet und für alle $i \in I$ diese jeweils durch die Permutation $\pi^{(i)}$ wie folgt geordnet

$$\frac{|\hat{C}_{i,\pi_1^{(i)}}|}{\left\| \hat{C}_{:, \pi_1^{(i)}} \right\|_\infty} \leq \frac{|\hat{C}_{i,\pi_2^{(i)}}|}{\left\| \hat{C}_{:, \pi_2^{(i)}} \right\|_\infty} \leq \dots \leq \frac{|\hat{C}_{i,\pi_s^{(i)}}|}{\left\| \hat{C}_{:, \pi_s^{(i)}} \right\|_\infty}.$$

Unkorrelierte Faktoren

Für das Ziel möglichst unkorrelierter Faktoren ist die direkte Anwendung der Stichproben-Kovarianz der Teilfaktoren in \hat{A} ungeeignet. Die Einarbeitung in eine Quadratmitteloptimierung führt auf die Minimierung von $\text{Cov}(\hat{A}_{i,:}, \hat{A}_{j,:})^2$ und nicht von $\text{Cov}(\hat{A}_{i,:}, \hat{A}_{j,:})$, was für durchaus angestrebte negative Werte von $\text{Cov}(\hat{A}_{i,:}, \hat{A}_{j,:})$ nicht sinnvoll ist. Alternativen sind

$$\tilde{f}_{12}(\hat{A}) = \sum_{i=1}^{s-1} \sum_{j=i+1}^s \exp(\alpha \text{Cov}(\hat{A}_{i,:}, \hat{A}_{j,:}))$$

mit $\alpha > 0$ oder (mit dem Hintergrund nichtnegativer Einträge in \hat{A}) die Summe der paarweisen Skalarprodukte

$$f_{12} : \mathbb{R}^{s \times n} \rightarrow \mathbb{R}, \quad f_{12}(\hat{A}) = \sum_{i=1}^{s-1} \sum_{j=i+1}^s \hat{A}_{i,:} \hat{A}_{j,:}^T$$

In der Anwendung hat sich dabei f_{12} als geeignet herausgestellt.

An einem ähnlichen Punkt setzt

$$f_{13} : \mathbb{R}^{k \times s} \times \mathbb{R}^{s \times n} \rightarrow \mathbb{R}, \quad f_{13}(\hat{C}, \hat{A}) = \left(\sigma_s(\hat{C}) + \epsilon_c \right)^{-2} + \left(\sigma_s(\hat{A}) + \epsilon_a \right)^{-2}$$

an, welche mit passend gewählten $\epsilon_c, \epsilon_a > 0$ einer „starken linearen Abhängigkeit“ unter den Faktoren jeweils in \hat{C} und \hat{A} entgegenwirkt.

Transformation der Straf- und der Regularisierungsfunktionen und die zusammengesetzte Zielfunktion

Durch die Ansätze aus (6.2) sind die Funktionen $f_i(\hat{C}, \hat{A})$, $f_i(\hat{C})$ beziehungsweise $f_i(\hat{A})$, $i = 0, \dots, 13$, indirekt Funktionen in Abhängigkeit von T

$$\begin{aligned} f_i(\hat{C}, \hat{A}) &= f_i(U\Sigma_{:,1:z}T^+, TV_{:,1:z}^T), & i &= 0, 13, \\ f_i(\hat{C}) &= f_i(U\Sigma_{:,1:z}T^+), & i &= 3, 7, 9, 11, \\ f_i(\hat{A}) &= f_i(TV_{:,1:z}^T), & i &= 1, 2, 4, 5, 6, 8, 10, 12. \end{aligned}$$

Von dieser Stelle an werden sie auch als solche betrachtet

$$\begin{aligned} \bar{f}_i &: \mathbb{R}^{s \times z} \rightarrow \mathbb{R}, \\ \bar{f}_i(T) &= f_i(U\Sigma_{:,1:z}T^+, TV_{:,1:z}^T), & i &= 0, 13, \\ \bar{f}_i(T) &= f_i(U\Sigma_{:,1:z}T^+), & i &= 3, 7, 9, 11, \\ \bar{f}_i(T) &= f_i(TV_{:,1:z}^T), & i &= 1, 2, 4, 5, 6, 8, 10, 12. \end{aligned}$$

Die aus der regularisierten Faktorisierungsaufgabe abgeleitete und aus den einzelnen Teilfunktionen zusammengesetzte Gesamtzielfunktion lautet demnach $F : \mathbb{R}^{s \times z} \rightarrow \mathbb{R}$,

$$F(T) = \frac{1}{2} \left(\bar{f}_0(T) + \sum_{i=1}^{13} \gamma_i^2 \bar{f}_i(T) \right) \quad (6.11)$$

mit

$$F(T) \rightarrow \min$$

als Ziel.

Die Modifikation der Abstandsfunktion

Die Abstandsfunktion f_0 lässt sich wie folgt modifizieren. Da bereits eine Singulärwertzerlegung der Matrix A durch den Ansatz aus (6.2) bestimmt wird, ist es aufgrund des verminderten Rechenaufwandes sinnvoll

$$\bar{f}_0(T) = \|\Sigma_{1:z,1:z}(I_z - T^+T)\|_F^2 + \sum_{i=z+1}^{\min\{k,n\}} \sigma_i^2$$

zu verwenden. Für eine lokale Niedrigrangapproximation auf dem Indexbereich I ergibt sich

$$\bar{f}_0(T) = \|\Sigma_{1:z,1:z}(I_z - T^+T)V_{I,1:z}^T\|_F^2 + \|\text{diag}(\sigma_{z+1}, \dots, \sigma_{\min\{k,n\}})V_{I,z+1:\min\{k,n\}}^T\|_F^2.$$

Stetigkeits- und Differenzierbarkeitsbetrachtungen

Für die Lösung des Optimierungsproblems aus (6.1) stellt sich die Frage nach der Stetigkeit der einzelnen Straf- und Regularisierungsfunktionen und dazu zunächst nach der Stetigkeit bei der Bildung der Pseudoinversen.

Bemerkung 6.5. Die Abbildung $\phi : \mathbb{R}^{s \times z} \rightarrow \mathbb{R}^{z \times s}$, die jeder Matrix T ihre Pseudoinverse zuordnet, ist nicht stetig. Sei $T = (\varepsilon)$, dann ergibt sich dazu $T^+ = (1/\varepsilon)$ und es gilt für $\varepsilon \rightarrow 0$

$$T \rightarrow \bar{T} = (0).$$

Demgegenüber steht jedoch

$$T^+ = (1/\varepsilon) \rightarrow (0) = \overline{T}^+.$$

Unter geeigneten Voraussetzungen lässt sich die Stetigkeit jedoch sicherstellen, wie das folgende Lemma von Rakocëvić zeigt.

Lemma 6.6. (Rakocëvić, [73]) Seien $A \in \mathbb{R}^{s \times z}$ sowie $\{A_n\}_{n \in \mathbb{N}}$ eine Folge mit ebenfalls $A_n \in \mathbb{R}^{s \times z}$ für $n \in \mathbb{N}$. Weiter gelte $A_n \rightarrow A$.

Dann gilt

$$A_n^+ \rightarrow A^+ \Leftrightarrow \exists n_0 > 0 : \text{rank}(A_n) = \text{rank}(A), \quad \forall n \geq n_0.$$

Über die Stetigkeit der einzelnen Zielfunktionen lässt sich Folgendes aussagen.

Bemerkung 6.7. Die Funktionen f_i , $i \in \{i \in \mathbb{N} : i \leq 12\} \setminus \{11\}$, sind stetig differenzierbar, wobei für f_9 beziehungsweise f_{10} der Zusatz $\|\hat{C}_{:,i}\| > 0$ beziehungsweise $\|\hat{A}_{i,:}\| > 0$ für $i = 1, \dots, s$ notwendig ist. Für zwei Matrizen $\hat{C} \in \mathbb{R}^{k \times s}$, $\hat{A} \in \mathbb{R}^{s \times n}$ mit $\sigma_s(\hat{C}) > \epsilon_c$, $\sigma_s(\hat{A}) > \epsilon_a$ ist auch f_{13} stetig differenzierbar (siehe dazu beispielsweise [78], Satz 1). Die Funktionen f_{11}, f_{13} sind stetig.

Für eine Matrix $T \in \mathbb{R}^{s \times z}$ mit vollem Rang s sowie $\gamma_{11} = 0$, $\sigma_s(\Sigma_{1:,1:,z} T^+) > \epsilon_c$, $\sigma_s(T) > \epsilon_a$ ist $F(T)$ aus (6.11) nach Lemma 6.6 und Bemerkung 6.7 stetig differenzierbar, da gilt $\text{rank}(U_{:,1:,z}) = \text{rank}(V_{:,1:,z}) = z$.

Der Rechenaufwand

Da es sich beim Berechnen der regularisierten nichtnegativen Niedrigrangapproximation um Relationen $\min\{k, n\} \gg z \geq s$ handelt, sind bei einer Auflistung aller notwendigen Multiplikationen die Dimensionen der Ausgangsmatrix die führenden Größen für eine Funktionsauswertung $F(T)$ und der Aufwand entspricht

$$\text{ROW} = \mathcal{O}(k + n).$$

Die Aufspaltung in Komponenten zur Quadratmittelloptimierung

Um bei der Optimierung der Zielfunktion aus (6.1) eine bessere Aufschlüsselung der einzelnen Komponenten zu erhalten, ist es sinnvoll, die Auswertung der Frobeniusnorm der einzelnen Teilfunktionen, so weit möglich, komponentenweise in einen Algorithmus zur Quadratmittelloptimierung einzubinden. Es lassen sich so die Informationen über die partiellen Ableitungen zur Bestimmung des Gradienten durch finite Differenzen numerisch präziser verarbeiten. Aus nomenklaturtechnischen Gründen ist \bar{f}_0 davon ausgenommen, wird jedoch durch $\bar{\bar{f}}_0(T) = \sqrt{\bar{f}_0(T)}$ ersetzt. Insgesamt ergibt sich somit für p Straf- und Regularisierungsquadratkomponenten

$$F(T) = \frac{1}{2} \left(\bar{\bar{f}}_0^2(T) + \sum_{i=1}^p \gamma_i^2 \bar{\bar{f}}_i^2(T) \right). \quad (6.12)$$

6.2 Dynamik eines Matrixfaktors und kinetische Analyse

Für spezielle den Daten zugrunde liegende Modellierungen gibt es eine weitere Regularisierungsmöglichkeit. Sind die einzelnen Teilkomponenten aus einem der Faktoren über eine Differentialgleichung untereinander verknüpft, so lässt sich die Lösungsvielfalt weiter einschränken.

6.2.1 Zusammenhang der Teilfaktoren über ein Differentialgleichungssystem

In Abschnitt 2.3.1 ist die Anwendung von nichtnegativen Matrixfaktorisierungen auf spektroskopische Daten erläutert. Sind die Einträge in A spektroskopische Daten einer chemischen Reaktion, so beschreibt der Faktor \hat{C} die Konzentrationsverläufe der teilnehmenden unabhängigen Komponenten. Diese Konzentrationsprofile stehen über eine gewöhnliche Differentialgleichung in Zusammenhang. Ist diese dabei konkret gegeben, so lässt sich der Faktor \hat{C} anhand der Eduktkonzentrationen bestimmen und dazu unmittelbar \hat{A} berechnen.

Enthält die Differentialgleichung jedoch unbekannte Parameter, so lässt sich deren Berechnung mit der Bestimmung der Faktorisierung koppeln. In Abhängigkeit des Parameters ϕ und des Faktors \hat{C} wird ein Anfangswertproblem gelöst und die Auswertung dieser Lösung an den Stützstellen $\bar{t} \in \mathbb{R}^k$ mit \hat{C} abgeglichen.

Kopplung von Faktorisierung und Parametrisierung in einem Ausgleichsproblem

Sei für den Faktor $c(t)$ die Differentialgleichung

$$\dot{c}(t) = f(t, c; \phi)$$

mit $f : \mathbb{R} \times \mathbb{R}^s \times \mathbb{R}^\nu \rightarrow \mathbb{R}^s$ und den Parametern ϕ_i , $i = 1, \dots, \nu$, gegeben. Zur Anpassung ist eine korrekte Skalierung der Werte in \hat{C} notwendig. Einige Möglichkeiten den Skalierungsfaktor $\alpha \in \mathbb{R}^s$ für die einzelnen Spalten zu berechnen, werden in Abschnitt 6.2.3 erläutert. Die absoluten Faktoren seien durch $C^{(\text{abs})} \in \mathbb{R}^{k \times s}$ gegeben.

Aus der Differentialgleichung mit dem Parameter ϕ , dem Intervall $t \in [\bar{t}_1, \bar{t}_k]$ und der ersten Zeile von $C^{(\text{abs})}$ setzt sich ein Anfangswertproblem zusammen. Ziel ist es, die Auswertung dessen Lösung $c(t)$ an den Stützstellen \bar{t}_i , $i = 1, \dots, k$, und die Einträge in $C^{(\text{abs})}$ einander anzupassen. Es liegt das Ausgleichsproblem

$$\sum_{i=2}^k \sum_{j=1}^s (C_{i,j}^{(\text{dgl})}(\phi) - C_{i,j}^{(\text{abs})})^2 \rightarrow \min \quad (6.13)$$

vor, wobei die Einträge in $C^{(\text{dgl})}(\phi)$ die Auswertungen der Lösung $c(t)$ an den Stützstellen \bar{t} sind.

Folge von Anfangswertproblemen

Um die Matrix $C^{(\text{dgl})}$ zeilenweise zu berechnen lässt sich das zu lösende Anfangswertproblem in eine Folge aus $k - 1$ Anfangswertproblemen zerlegen. Das erste lautet

$$\begin{aligned} \dot{c}(t) &= f(t, c; \phi), & t &\in [\bar{t}_1, \bar{t}_2], \\ c(\bar{t}_1) &= C_{1,:}^{(\text{dgl})} \end{aligned}$$

und der Vergleichswert in (6.13) ist $C_{2,:}^{(\text{dgl})} = c(\bar{t}_2)$.

Die Fortsetzung der Folge von Anfangswertproblemen ist auf zwei verschiedene Weisen möglich. Einerseits lassen sich die berechneten Werte $C_{2,:}^{(\text{dgl})}$ als Anfangswerte für das zweite Anfangswertproblem

$$\begin{aligned} \dot{c}(t) &= f(t, c; \phi), & t &\in [\bar{t}_2, \bar{t}_3], \\ c(\bar{t}_2) &= C_{2,:}^{(\text{dgl})} \end{aligned}$$

nutzen (induktive Folge von Anfangswertproblemen). Andererseits ist es auch möglich, eine Lösung mit den Anfangswerten $C_{2,:}^{(\text{abs})}$ (autonome Folge) zu berechnen

$$\begin{aligned}\dot{c}(t) &= f(t, c; \phi), & t \in [\bar{t}_2, \bar{t}_3], \\ c(\bar{t}_2) &= C_{2,:}^{(\text{abs})}.\end{aligned}$$

Die Auswertung von $c(t)$ am Intervallende führt jeweils auf die Anpassungswerte $C_{3,:}^{(\text{dgl})} = c(\bar{t}_3)$. Weiter liefert die jeweilige Fortsetzung dieser Vorgehensweisen (induktiv oder autonom) für die Zeitintervalle $[\bar{t}_i, \bar{t}_{i+1}]$, $i = 3, \dots, k-1$, eine Folge von Anfangswertproblemen. Die Auswertungen deren Lösungen an den Intervallenden sind die Einträge in $C^{(\text{dgl})}$ für die Ausgleichsrechnung.

6.2.2 Die Zielfunktion

Für die Zielfunktion ist zu beachten, dass $C^{(\text{abs})}$ in (6.13) von \hat{C} abhängt. Mit den anzupassenden Parametern ϕ ergibt sich insgesamt die Zielfunktion

$$f_{14} : \mathbb{R}^{k \times s} \times \mathbb{R}^\nu \rightarrow \mathbb{R}, \quad f_{14}(\hat{C}, \phi) = \sum_{i=2}^k \sum_{j=1}^s (C_{i,j}^{(\text{dgl})}(\phi, \hat{C}) - C_{i,j}^{(\text{abs})}(\hat{C}))^2.$$

6.2.3 Skalierung der Faktoren

Ein Problem ist die Skalierung der Spalten des Faktors \hat{C} , welche für die korrekte Lösung des Anfangswertproblems notwendig ist. Nun folgend sind verschiedene Möglichkeiten genannt, diese zu bestimmen.

1. Sind Stützstellen \bar{t}_{i_0} und Werte $c_i(\bar{t}_{i_0}) > 0$ für alle $i = 1, \dots, s$ bekannt, so erfolgt die Anpassung darauf.
2. Ist zu den Profilen ein Wert $c^{(\text{ges})} > 0$ und ein Zusammenhang der Form

$$c^{(\text{ges})} = \sum_{i=1}^s \beta_i c_i(t), \quad \forall t \geq 0, \quad (6.14)$$

mit $\beta_i > 0$, $i = 1, \dots, s$, bekannt, so ergeben sich $C_{:,i}^{(\text{abs})} = \alpha_i \hat{C}_{:,i}$, $i = 1, \dots, s$, mit α als Lösung des restringierten linearen Ausgleichsproblems

$$\min_{\alpha \in \mathbb{R}_+^s} \left\| (c^{(\text{ges})}, \dots, c^{(\text{ges})})^T - \hat{C} \text{diag}(\beta) \alpha \right\|_2^2.$$

3. Sind ein Index i_0 , eine Stützstelle t_0 und ein Wert $c_{i_0}(t_0) > 0$ bekannt sowie ein Zusammenhang wie in (6.14), jedoch mit unbekanntem $c^{(\text{ges})}$, so lassen sich die weiteren Profile $c_i(t)$, $i \in \{1, \dots, s\} \setminus \{i_0\}$ über ein lineares Ausgleichsproblem bestimmen. Seien dazu $i_0 = s$ und $C_{:,s}^{(\text{abs})}$ die bereits skalierte Linie zu $\hat{C}_{:,s}$, dann ergeben sich $C_{:,i}^{(\text{abs})} = \alpha_i \hat{C}_{:,i}$ für $i = 1, \dots, s-1$ mit α als Lösung des restringierten linearen Ausgleichsproblems

$$\min_{\alpha \in \mathbb{R}_+^{s-1}} \sum_{i=1}^k \left(\beta_s \left(C_{i,s}^{(\text{abs})} - \frac{1}{k} \sum_{l=1}^k C_{l,s}^{(\text{abs})} \right) + \sum_{j=1}^{s-1} \alpha_j \beta_j \left(\hat{C}_{i,j} - \frac{1}{k} \sum_{l=1}^k \hat{C}_{l,j} \right) \right)^2. \quad (6.15)$$

Die Koeffizientenmatrix des Ausgleichsproblems ist dabei durch $\text{rank}(\hat{C}_{:,1:s-1}) = s-1$ ebenfalls von vollem Rang $s-1$, sofern gilt $(1, \dots, 1)^T \notin \text{im}(\hat{C}_{:,1:s-1})$.

4. Sind für die Punkte 2 oder 3 zu mehreren, aber nicht allen Profilen $c_i(t)$, $i = 1, \dots, s$, konkrete Werte größer Null bekannt, so lassen sich die Ausgleichsprobleme diesbezüglich modifizieren.

6.2.4 Einschränkung der Modellfreiheit

Sind die Regularisierungen aus den Funktionen f_i , $i = 4, \dots, 13$, bereits einschneidende Eingriffe in die Modellfreiheit der Niedrigrangapproximation, so ist die Kopplung eines der Faktoren mit einer zu parametrisierenden Differentialgleichung eine weitere starke Einschränkung aller möglichen Äquivalenzklassen an Lösungen. Handelt es sich bei dem kinetischen Modell nur um eine Vermutung, so lässt sich diese zwar möglicherweise ausschließen, im Normalfall aber nicht sicher als die einzige Möglichkeit belegen. Mit der Gewichtung in der aus allen Regularisierungen zusammengesetzten Zielfunktion aus (6.1) sollte dies dementsprechend berücksichtigt werden.

6.3 Methoden der nichtlinearen Ausgleichsrechnung

In diesem Abschnitt wird die Minimierung der Funktion aus (6.11) betrachtet, welche bei der Umsetzung der vorangegangenen Theorie ein Kernstück der Selbstmodellierung darstellt. Dazu wird das Gauß-Newton Verfahren eingeführt und anschließend daran die ACM FORTRAN Bibliothek NL2SOL als leistungsstarker Optimierer vorgestellt. Für die Erläuterungen zum Gauß-Newton Verfahren sowie allgemein zur nichtlinearen Ausgleichsrechnung ist zudem auf Björck, [5], Nocedal und Wright, [65], sowie Kosmol, [49], verwiesen.

Um Inkonsistenzen mit den Dimensionen zu vermeiden, werden in diesem Kapitel nicht speziell die Funktionen f_i aus den vorherigen Teilabschnitten behandelt, sondern allgemein eine Quadratmittelloptimierung von Funktionen g_i . Weiter wird in diesem Abschnitt die Optimierung über einen Vektor $x \in \mathbb{R}^q$ und nicht über eine Matrix T betrachtet. Diese nomenklaturtechnische Modifikation hat keinerlei Einfluss auf die grundsätzliche Theorie.

Die unrestringierte Quadratmittelaufgabe

Das vorliegende nichtlineare Ausgleichsproblem zu $g : \mathbb{R}^q \rightarrow \mathbb{R}^p$ mit $p \geq q$ lautet $G : \mathbb{R}^q \rightarrow \mathbb{R}$,

$$G(x) = \frac{1}{2} \sum_{i=1}^p g_i^2(x), \quad G(x) \rightarrow \min. \quad (6.16)$$

Eigenschaften der Quadratmittelfunktion

Von der Funktion G ist zu erwarten, dass sie nicht konvex aber mindestens einmal stetig differenzierbar ist. Für die Zielfunktion aus (6.11) wäre dies im Normalfall für die Wahl $\gamma_{11} = \gamma_{13} = 0$ gegeben. Wie die folgende Bemerkung zeigt, ist F aus (6.11) zudem koerziv.

Bemerkung 6.8. *In dem vorliegenden Fall der regularisierten Niedrigrangapproximation ist die Zielfunktion F aus (6.11) koerziv, da aufgrund der Orthogonalität von V die stets verwendete Normierungsfunktion \bar{f}_1 koerziv ist. Weiter ist beispielsweise \bar{f}_6 ebenfalls koerziv.*

Nachweis: Sei $\{T^{(i)}\}_{i \in \mathbb{N}}$ eine Folge von Matrizen $T^{(i)} \in \mathbb{R}^{s \times z}$, $i = 1, 2, \dots$, mit

$$\lim_{i \rightarrow \infty} \|T^{(i)}\| = \infty.$$

Dann gilt

$$\begin{aligned} \bar{f}_1(T) &= \sum_{i=1}^s (1 - \|T_{i,:} V_{:,1:z}^T\|_\infty)^2 = \sum_{i=1}^s 1 - 2\|T_{i,:} V_{:,1:z}^T\|_\infty + \|T_{i,:} V_{:,1:z}^T\|_\infty^2 \\ &\geq \sum_{i=1}^s 1 - 2\|T_{i,:} V_{:,1:z}^T\|_2 + \frac{1}{n} \|T_{i,:} V_{:,1:z}^T\|_2^2 = s - 2\left(\sum_{i=1}^s \|T_{i,:}\|_2\right) + \frac{1}{n} \|T\|_F^2 \end{aligned}$$

und mit $\|x\|_1 \leq \sqrt{s}\|x\|_2$, für $x \in \mathbb{R}^s$, folgt

$$\bar{f}_1(T) \geq s - 2\sqrt{s}\|T\|_F + \frac{1}{n}\|T\|_F^2 \xrightarrow{i \rightarrow \infty} \infty.$$

Somit ist $\bar{f}_1(T)$ koerziv.

Die Funktion \bar{f}_6 ist durch ihre Definition über die von der Matrix $\text{diag}(b^{(6)})$ induzierten Norm ebenfalls koerziv, da zusätzlich $\text{rank}(V_{:,1:z}) = z$ gilt. \square

Minimalstelle der Quadratmittelfunktion

Eine Minimalstelle x^* des allgemeinen Optimierungsproblems aus (6.16) mit zweimal stetig differenzierbaren Funktionen $g_i(x)$, $i = 1, \dots, p$, ist charakterisiert durch die notwendigen Bedingungen

$$\begin{aligned} \nabla G(x^*) &= (J_g^T g)(x^*) = 0 & (\Leftrightarrow g(x^*) \in \ker(J_g(x^*)) \Leftrightarrow g(x^*) \perp J_g^T(x^*)), \\ H_G(x^*) &= \left(J_g^T J_g + \sum_{i=1}^p g_i H_{g_i} \right)(x^*) & \text{ist symmetrisch positiv semidefinit.} \end{aligned} \quad (6.17)$$

Eine hinreichende Bedingung für eine lokale Minimalstelle ist die Forderung, dass $H_G(x^*)$ positiv definit ist. Dabei ist $J_g \in \mathbb{R}^{p \times q}$ die Jacobimatrix der partiellen Ableitungen von g und die $H_{g_i} \in \mathbb{R}^{q \times q}$, $i = 1, \dots, p$, sind die Hessematrizen der zweiten partiellen Ableitungen der Funktionen g_i

$$\begin{aligned} J_g &= \left(\frac{\partial g_i}{\partial x_j} \right)_{i=1, \dots, p, j=1, \dots, q}, \\ H_{g_\ell} &= \left(\frac{\partial^2 g_\ell}{\partial x_i \partial x_j} \right)_{i,j=1, \dots, q}, \quad \ell = 1, \dots, p. \end{aligned}$$

Zum Bestimmen einer solchen Minimalstelle x^* lassen sich das Gauß-Newton Verfahren, [5, 49], Quasi-Newton Verfahren, [5, 49], oder modifizierte Verfahren dieser Klassen, [17, 19, 49], einsetzen.

6.3.1 Das Gauß-Newton Verfahren

Das Gauß-Newton Verfahren ist eine iterative Methode zur Lösung von nichtlinearen Ausgleichs-problemen für Funktionen $g : \mathbb{R}^q \rightarrow \mathbb{R}^p$. Das Ziel ist die Konstruktion einer Folge von Iterierten $x_k \in \mathbb{R}^q$, $k = 1, 2, \dots$, zum Startwert x_0 mit

$$\lim_{k \rightarrow \infty} \frac{1}{2} \sum_{i=1}^p g_i(x_k)^2 = \min_{x \in \mathbb{R}^q} \frac{1}{2} \sum_{i=1}^p g_i(x)^2.$$

Approximation durch ein lineares Modell

Sei x_k die aktuelle Iterierte. Um x_{k+1} zu bestimmen, wird die Taylor-Reihe von $g(x_k)$

$$g(x_k + \Delta_x) = g(x_k) + J_g(x_k)\Delta_x + \mathcal{O}((\|\Delta_x\|^2, \dots, \|\Delta_x\|^2)^T)$$

nach dem linearen Term abgeschnitten und die Funktion $g^{(k)} : \mathbb{R}^q \rightarrow \mathbb{R}^p$,

$$g^{(k)}(x) = g(x_k) + J_g(x_k)(x - x_k)$$

extrahiert. Für diese soll das lineare Ausgleichsproblem in x

$$G_k(x) = \frac{1}{2} \|g^{(k)}(x)\|_2^2 \rightarrow \min \quad (6.18)$$

gelöst werden. Die notwendige Bedingung für ein Optimum x_k^* dieser konvexen Funktion ist gleichzeitig hinreichend und lautet

$$\|\nabla G_k(x_k^*)\| = \|J_g^T(x_k)J_g(x_k)(x_k^* - x_k) + J_g^T(x_k)g(x_k)\| = 0. \quad (6.19)$$

Lösung des linearen Ausgleichsproblems

Die Lösung x_k^* von (6.18) liefert die neue Iterierte zur Optimierung der Funktion G

$$x_{k+1} = x_k^*$$

beziehungsweise

$$x_{k+1} = x_k + \lambda_k(x_k^* - x_k)$$

mit einer geeigneten Schrittweite λ_k bei Anwendung einer Schrittweitensteuerung. Die Lösung des linearen Ausgleichsproblems (6.18) in x , lässt sich durch Lösung des linearen Gleichungssystems in x_k^* aus (6.19) erreichen. Aus Stabilitätsgründen ist es jedoch sinnvoller das Ausgleichsproblem direkt mittels QR -Faktorisierung oder Singulärwertzerlegung zu lösen, vergleiche [62].

Das lineare Ausgleichsproblem aus (6.18) ist eindeutig lösbar, sofern dessen Matrix $J_g(x_k)$ den vollen Spaltenrang besitzt. Dann gilt

$$\|J_g(x_k)x + g(x_k) - J_g(x_k)x_k\|_2 = \|\Sigma V^T x - U^T(J_g(x_k)x_k - g(x_k))\|_2$$

mit einer Singulärwertzerlegung $J_g(x_k) = U\Sigma V^T$ und damit für x weiter

$$\begin{aligned} V^T x &= \Sigma^+ U^T (J_g(x_k)x_k - g(x_k)), \\ x &= V\Sigma^+ U^T (J_g(x_k)x_k - g(x_k)) = (J_g(x_k))^+ (J_g(x_k)x_k - g(x_k)). \end{aligned}$$

Hat $J_g(x_k)$ nicht den vollen Spaltenrang, so ergibt sich zunächst erneut

$$\Sigma V^T x = U^T (J_g(x_k)x_k - g(x_k)), \quad \text{rank}(\Sigma) = r < q.$$

Weiter folgt daraus die Forderung

$$\begin{aligned} (\Sigma V^T x)_{1:r} &= (U^T (J_g(x_k)x_k - g(x_k)))_{1:r}, \\ \Sigma_{1:r,1:r} V_{:,1:r}^T x &= (U^T (J_g(x_k)x_k - g(x_k)))_{1:r}, \end{aligned}$$

welche durch

$$x = (J_g(x_k))^+ (J_g(x_k)x_k - g(x_k))$$

erfüllt wird. Jedoch ist auch jeder Vektor der Form

$$x_0 = (J_g(x_k))^+ (J_g(x_k)x_k - g(x_k)) + V_{:,r+1:q}^T y, \quad y \in \mathbb{R}^{q-r},$$

Lösung des Ausgleichsproblems. Unter diesen wird diejenige Lösung mit der kleinsten euklidischen Norm weiter verwendet. Da gilt

$$(J_g(x_k))^+ (J_g(x_k)x_k - g(x_k)) \perp V_{:,r+1:q}^T,$$

lässt sich die Summe für eine Normberechnung von x aufspalten zu

$$\|x_0\|_2 = \|(J_g(x_k))^+ (J_g(x_k)x_k - g(x_k))\|_2 + \|V_{:,r+1:q}^T y\|_2$$

und es folgt die neue Iterierte

$$x = (J_g(x_k))^+ (J_g(x_k)x_k - g(x_k)).$$

Damit stimmt die gesuchte Lösung in beiden Fällen, $\text{rank}(J_g(x_k)) = q$ und $\text{rank}(J_g(x_k)) < q$, überein

$$x_k^* = (J_g(x_k))^+ (J_g(x_k)x_k - g(x_k)).$$

Abbruchbedingung

Als Abbruchkriterien für die Iteration des Gauß-Newton Verfahrens lassen sich die absolute Abbruchschranke $G(x_k) < \varepsilon_a$, die Norm des Gradienten

$$\|\nabla G(x_k)\| = \|(J_g^T g)(x_k)\| < \varepsilon$$

oder auch die x -Konvergenz

$$\|x_k - x_{k-1}\| < \epsilon$$

mit geeigneten $\varepsilon_a, \varepsilon, \epsilon > 0$ anwenden. Für weitere Abbruchkriterien ist auf [17, 49] verwiesen.

Konvergenz des Gauß-Newton Verfahrens

Zu Aussagen bezüglich der Konvergenz des Gauß-Newton Verfahrens lässt sich folgender Satz aus [49] anführen.

Satz 6.9. (Kosmol, [49]) Sei $g \in C^1(\mathbb{R}^q, \mathbb{R}^p)$ und sei J_g in $x^* \in \mathbb{R}^p$ mit der Konstanten $L > 0$ Lipschitz-stetig. Es gelte $\nabla G(x^*) = 0$, $\text{rank}(J_g(x^*)) = q$ und

$$L\|g(x^*)\| < \left\| (J_g(x^*)^T J_g(x^*))^{-1} \right\|^{-1}.$$

Dann gibt es eine Umgebung K von x^* , so dass das Gauß-Newton Verfahren für jeden Startpunkt aus K durchführbar ist, die gesamte Iterationsfolge bleibt in K und konvergiert mindestens Q -linear gegen x^* . Ist zusätzlich $g(x^*) = 0$, so ist die Konvergenz mindestens Q -quadratisch.

Definition 6.10. Eine Folge $(x_k)_{k \in \mathbb{N}}$ heißt Q -linear, beziehungsweise Q -quadratisch konvergent gegen x^* , falls gilt

1. $(x_k)_{k \in \mathbb{N}}$ konvergiert gegen x^* und
2. $\|x_{k+1} - x^*\| \leq C\|x_k - x^*\|$ mit einem $C \in [0, 1)$, beziehungsweise $\|x_{k+1} - x^*\| \leq C\|x_k - x^*\|^2$ mit einem $C \in [0, 1)$.

Bemerkung 6.11. Der Fall $G(x^*) = 0$ ist für die regularisierte nichtnegative Niedrigrangapproximation unter Verwendung der zusätzlichen Regularisierungen f_i , $i = 4, 5, 6, 7, 13$, nicht gegeben. Für die idealisierte Faktorisierungsaufgabe 2.3 ist $G(x^*) = 0$ bei ausschließlicher Verwendung von f_i , $i = 0, 1, 2, 3$, erfüllt.

Bemerkung 6.12. Die Lösung des linearen Ausgleichproblems aus (6.18) liefert eine Minimierung der quadratischen Funktion

$$G_k(x) = \frac{1}{2}g^T(x_k)g(x_k) + (x - x_k)^T J_g(x_k)^T g(x_k) + \frac{1}{2}(x - x_k)^T J_g^T(x_k)J_g(x_k)(x - x_k). \quad (6.20)$$

Ein Vergleich mit der Taylorentwicklung der Quadratmittelfunktion G im Punkt x_k bis zur 2. Ordnung ergibt den Defekt

$$G_k(x) - G(x) \approx \frac{1}{2}(x - x_k)^T \left(\sum_{i=1}^p g_i(x_k)H_{g_i}(x_k) \right) (x - x_k). \quad (6.21)$$

Dieser hängt offenbar direkt von den Funktionswerten $g_i(x_k)$, $i = 1, \dots, p$, ab. Wie in Bemerkung 6.11 angemerkt, verschwinden die Funktionswerte $g_i(x)$ im Normalfall nicht einmal in der gesuchten Minimalstelle x^* und der Defekt sollte demzufolge nicht vernachlässigt werden.

Da das Optimierungsproblem der regularisierten Niedrigrangapproximation durch viele Zwischenminima charakterisiert ist, wird zu dessen Lösung nicht das Gauß-Newton Verfahren angewendet, sondern eine Weiterentwicklung. Ein solches leistungsstarkes Optimierungsprogramm ist NL2SOL.

6.3.2 Das Programmpaket NL2SOL

Das Programmpaket NL2SOL ist eine ACM FORTRAN Bibliothek zur Lösung von nichtlinearen Ausgleichsproblemen. Der Algorithmus basiert auf einer Kombination verschiedener Methoden, die durch geschickte Verknüpfung die große Leistungsfähigkeit des Verfahrens ausmachen. Zum Spektrum der Methoden zählen das klassische und ein verbessertes Gauß-Newton Verfahren, Quasi-Newton Methoden und ein *trust-region* Verfahren. Für tiefgreifendere Erläuterungen ist auf [17] verwiesen.

Korrekturterm zum Gauß-Newton Verfahren

Entscheidender Ansatzpunkt für eine Verbesserung des Gauß-Newton Verfahrens ist die Bemerkung 6.12. Um die sehr aufwendige Bestimmung der Hessematrizen der Einzelfunktionen (Aufwand: $\sim q^2$ zusätzliche Funktionsaufrufe) zu vermeiden, wird eine Folge von Näherungen dieser konstruiert, ähnlich wie bei Quasi-Newton Methoden. Zur Konstruktion einer Näherung

$$S_{k+1} \approx \sum_{i=1}^p g_i(x_{k+1})H_{g_i}(x_{k+1})$$

werden die bekannten $J_{k+1} := J_g(x_{k+1})$, $g_{k+1} := g(x_{k+1})$ und $J_k := J_g(x_k)$, $g_k := g(x_k)$ eingesetzt. Dabei ergibt sich mit $\Delta x_k = x_{k+1} - x_k$ die Forderung

$$\begin{aligned} S_{k+1}\Delta x_k &= \sum_{i=1}^p g_i(x_{k+1})H_{g_i}(x_{k+1})\Delta x_k \\ &\approx \sum_{i=1}^p g_i(x_{k+1})(\nabla g_i(x_{k+1}) - \nabla g_i(x_k)) \\ &= J_{k+1}^T g_{k+1} - J_k^T g_{k+1} =: y_k. \end{aligned} \quad (6.22)$$

Initialisiert wird die Folge mit $S_0 \in \mathbb{R}^{q \times q}$, $S_0 = 0$. Der erste Schritt zu x_1 ist also einer mit dem klassischen Gauß-Newton Verfahren. Weiter wird obigen Überlegungen zufolge S_{k+1} derart

gesucht, dass $S_{k+1} \in Q = \{S : S = S^T \text{ und } S\Delta x_k = y_k\}$, wobei y_k beispielsweise wie in (6.22) bestimmt wird oder auch nach Formeln von Broyden-Dennis oder Betts, vergleiche [17]. Als eine mögliche Korrekturformel für S ist

$$S_{k+1} = S_k + \frac{(y_k - S_k \Delta x_k)v^T + v(y_k - S_k \Delta x_k)^T}{\Delta x_k^T v} - \frac{\Delta x_k^T (y_k - S_k \Delta x_k)vv^T}{(\Delta x_k^T v)^2}$$

mit

$$v = \nabla g(x_{k+1}) - \nabla g(x_k) = J_{k+1}^T g_{k+1} - J_k^T g_k$$

in [17] angegeben.

Beschränkung auf eine Umgebung um die aktuelle Iterierte

Die Approximation durch das quadratische Gauß-Newton Modell aus (6.20) oder mit der Verbesserung um den Fehlerterm aus (6.21) sind nur lokal gültig. Speziell für die ersten Schritte, welche unter Umständen sehr groß sein können, lassen sich die lokalen quadratischen Modell nur schwer halten. Dies ist der Ausgangspunkt des *trust-region* Verfahrens; zu einer aktuellen Iterierten wird die neue Iterierte nur in einer vorgegebenen Umgebung um die aktuelle gesucht. In den einzelnen Iterationen erfolgt mit den hinzugewonnenen Informationen der neuen Punkte eine Anpassung der Größe und der Form der *trust-region*. Insgesamt wird in NL2SOL dazu ein Schritt der Form

$$\Delta x_k = -(H_k + \lambda_k D_k^2)^{-1} \nabla G(x_k) \quad (6.23)$$

angewendet. Die Matrix $H_k = J_g^T(x_k)J_g(x_k) + S_k$ ist dabei eine Approximation an die Hessematrix von G in x_k , vergleiche dazu (6.17), (6.21) und (6.22), D_k eine Diagonalmatrix zur Skalierung und $\lambda_k \geq 0$ eine weitere Schranke. Für weiterführende Erläuterungen ist hierbei direkt auf die zu NL2SOL gehörende Veröffentlichung [17] sowie allgemein zum *trust-region* Ansatz auf [5] verwiesen.

Die Abbruchbedingungen

Als letzten Punkt bezüglich NL2SOL werden nun verschiedene Abbruchkriterien angeführt. Das erste Abbruchkriterium ist die „absolute Funktionenkonvergenz“

$$g(x_k) < \varepsilon_A.$$

Dieses Kriterium behandelt gleichzeitig den sehr seltenen Fall mit $x^* = 0$ und $G(x^*) = 0$, welcher Probleme bei den späteren Kriterien der „ x -Konvergenz“ und der „relativen Funktionenkonvergenz“ erzeugen würde. Allen weiteren Abbruchbedingungen wird zunächst die Forderung einer nicht zu großen Fehlerreduktion

$$G(x_k) - G(x_k + \Delta x_k) \leq 2(G(x_k) - Q_k(x_k + \Delta x_k)) \quad (6.24)$$

auferlegt. Dabei ist Q_k das aktuelle quadratische Modell, für einen Gauß-Newton Schritt beispielsweise G_k aus (6.20). Gilt andernfalls die Ungleichung (6.24) nicht, so kann eine weitere Verbesserung, als es das aktuelle quadratische Modell vermuten lässt, erwartet werden.

Das Konvergenzkriterium der „ x -Konvergenz“ setzt ein, falls von dem Abstand zum unbekannten tatsächlichen Minimum x^* ($H_g(x^*)$ positiv definit) erwartet werden kann, dass dieser kleiner als ε_x ist. Entschieden wird dies, falls für einen Newton Schritt ($\lambda_k = 0$ in (6.23))

$$\text{RELDX}(x_k, x_k + \Delta x_k, D_k) = \frac{\|D_k(\Delta x_k)\|_\infty}{\|D_k(|x_k + \Delta x_k| + |x_k|)\|_\infty} \leq \varepsilon_x$$

gilt mit der Skalierungsmatrix $D_k = \text{diag}(d_1^k, \dots, d_q^k)$ aus (6.23) und der komponentenweisen Auswertung von $|x_k + \Delta x_k|, |x_k|$.

Die „relative Funktionenkonvergenz“

$$\text{RELDF}(x_k) = \frac{G(x_k) - Q_k(x_k - H_K^{-1} \nabla G(x_k))}{G(x_k)} \leq \varepsilon_R$$

oder eine Kombination mit der „ x -Konvergenz“ komplettieren die erfolgreichen Abbruchkriterien.

Eine „singuläre Konvergenz“ wird erkannt, falls die Hessematrix H_G einen echten Kern besitzt oder so schlecht konditioniert ist, dass sie „fast singulär“ ist. Greift keines der bisher genannten Abbruchkriterien und wird (6.24) nicht erfüllt, so kann von einer „falschen Konvergenz“ gegen einen nichtkritischen Punkt ausgegangen werden, falls sich gleichzeitig dazu eine Stagnation in x -Richtung

$$\text{RELDX}(x_k, x_k + \Delta x_k, D_k) < \varepsilon_F$$

einstellt, wobei für gewöhnlich $\varepsilon_F < \varepsilon_x$ gilt. Ursachen dafür sind möglicherweise zu scharf gewählte Schranken $\varepsilon_A, \varepsilon_x, \varepsilon_R$ (im Vergleich zur Genauigkeit mit der G und J_G berechnet werden), dass J_G fehlerhaft ist oder dass G oder J_G in der Nähe von x_k eine Unstetigkeit besitzt.

6.4 Anwendung von Multi-Level-Methoden

Ein zentrales Problem bei der Berechnung einer regularisierten Niedrigrangapproximation mittels der Optimierung von F aus (6.11) ist die Existenz vieler lokaler Minimalstellen von F .

Um dem Problem der vielen lokalen aber nicht globalen Minimalstellen entgegenzutreten, wird in PCD die Strategie der mehrfachen Optimierung mit Zufallsstartvektoren zusammen mit einer Multi-Level-Strategie verfolgt.

6.4.1 Mehrfachoptimierung mit einer zufälligen Wahl von Startvektoren

Für eine Anzahl von m zufälligen Startmatrizen wird eine Optimierung durchgeführt. Es ergeben sich in der Regel

$$T^{(i)}, \quad i = 1, \dots, m,$$

als lokale Minimalstellen von $F(T)$ mit den zugehörigen Funktionswerten

$$F(T^{(i)}), \quad i = 1, \dots, m.$$

Aus den lokalen Minimalstellen wird dasjenige $T^{(i_0)}$ mit dem kleinsten Funktionswert ausgewählt,

$$F(T^{(i_0)}) \leq F(T^{(i)}), \quad i = 1, \dots, m. \quad (6.25)$$

Durch die uneindeutige Sortierung existieren für die aus dem restringierten Faktorisierungsproblem 2.5 abgeleitete Optimierungsaufgabe für (6.11) in der Regel (d.h. bei sinnvoller Wahl der Gewichtungen in γ) mindestens $s!$ verschiedene globale Minimalstellen. Eine Garantie, dass eine Optimierung auf eines dieser führt, ist nicht gegeben.

Diese Mehrfachoptimierung vervielfacht wiederum den Rechenaufwand, da der Hauptteil, abgesehen von der Singulärwertzerlegung, insgesamt m mal durchlaufen wird. An dieser Stelle gilt es einen geeigneten Kompromiss aus der Anzahl der Startmatrizen und einem vertretbaren Rechenaufwand zu finden. Die Zahl m der Zufallsstartmatrizen ist dabei über Anzahl und Lage der lokalen Minima problemabhängig.

6.4.2 Die Anwendung einer Dünngitterapproximation

Einen Ausweg aus dem deutlich erhöhtem Rechenaufwand bietet eine Ausdünnung der Daten. Der regularisierten Niedrigrangapproximation aus 2.5 liegt der idealisierte Produktansatz 2.2 mit den zwei Faktoren $\hat{c}, \hat{a} \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s)$ zugrunde. Für unterschiedliche Diskretisierungsgitter, soweit geeignet gewählt, ergeben sich dabei sehr ähnlich strukturierte Daten.

Dies eröffnet die Möglichkeit, von der Grunddiskretisierung $\bar{t} \in \mathbb{R}^k$, $\bar{x} \in \mathbb{R}^n$ und der zugehörigen Datenmatrix A ausgehend, eine Ausdünnung des Gitters vorzunehmen und darauf eine Niedrigrangapproximation zu bestimmen.

Das Ziel der Multi-Level-Strategie

Der Vorteil dieser Ausdünnung ist, dass sich die Optimierungsfunktion aus (6.11) zur Bestimmung der Niedrigrangapproximation vereinfacht und sich somit der Rechenaufwand für eine Funktionsauswertung reduziert. Damit lässt sich der Nachteil der Mehrfachoptimierung unter Umständen mehr als egalisieren. Aufgrund der Struktur der Daten ist diese Niedrigrangapproximation auf ein weniger stark ausgedünntes Gitter übertragbar und es ergibt sich ein geeigneter Startvektor. Somit wird die Mehrfachoptimierung unnötig. Für einen fallenden Grad der Ausdünnung lässt sich dies so weit anwenden, bis eine Niedrigrangapproximation der Ausgangsdaten vorliegt.

Diese Idee ist in eine *Multi-Level-Strategie* verpackt, vergleichbar mit den Mehrgittermethoden für lineare Gleichungssysteme von diskretisierten partiellen Differentialgleichungen.

Der Aufbau einer Dünngitterhierarchie

Zum Aufbau der Dünngitterhierarchie in verschiedenen Stufen wird Folgendes benötigt.

1. Die Grunddiskretisierung oder Wurzelstufe: Diskretisierung $\bar{t} \in \mathbb{R}^k$, $\bar{x} \in \mathbb{R}^n$ und Datenmatrix $A \in \mathbb{R}^{k \times n}$.
2. Eine monoton wachsende Folge P : Gewählt sei für diesen Abschnitt die geometrische Folge $P_i = 2^i$, $i = 0, 1, 2, \dots$.
3. Ausdünnung der Grunddiskretisierung bezüglich der Folge P mit den maximalen Grobgitterstufen $l_t, l_x \in \mathbb{N}$: Es resultieren für die in Punkt 2 gewählte Folge P die Grobgitter

$$\begin{aligned} t^{(i)} &\in \mathbb{R}^{\lfloor \frac{k}{2^i} \rfloor}, & t_{\nu}^{(i)} &= t_{1+(\nu-1)2^i}, & \nu &= 1, \dots, \left\lfloor \frac{k}{2^i} \right\rfloor, & i &= 0, \dots, l_t, \\ x^{(j)} &\in \mathbb{R}^{\lfloor \frac{n}{2^j} \rfloor}, & x_{\mu}^{(j)} &= x_{1+(\mu-1)2^j}, & \mu &= 1, \dots, \left\lfloor \frac{n}{2^j} \right\rfloor, & j &= 0, \dots, l_x. \end{aligned}$$

Für die Daten ergeben sich

$$\begin{aligned} A^{(i,j)} &\in \mathbb{R}^{\lfloor \frac{k}{2^i} \rfloor \times \lfloor \frac{n}{2^j} \rfloor}, & i &= 0, \dots, l_t, & j &= 0, \dots, l_x \\ A_{\nu,\mu}^{(i,j)} &= A_{1+(\nu-1)2^i, 1+(\mu-1)2^j}, & \nu &= 1, \dots, \left\lfloor \frac{k}{2^i} \right\rfloor, & \mu &= 1, \dots, \left\lfloor \frac{n}{2^j} \right\rfloor, \end{aligned}$$

und insgesamt resultieren zusammengehörige Tripel $(t^{(i)}, x^{(j)}, A^{(i,j)})$, $i = 0, \dots, l_t$, $j = 0, \dots, l_x$.

4. Auswahl der Ausdünnungskombination: Dazu seien zwei monoton nichtfallende Folgen τ und χ der Länge $\ell \in \mathbb{N}$ mit folgenden Eigenschaften definiert

$$\begin{aligned} \tau_1 &= 0, & \tau_\ell &= l_t, \\ \chi_1 &= 0, & \chi_\ell &= l_x, \\ \tau_{i+1} + \chi_{i+1} &\geq \tau_i + \chi_i + 1, & i &= 0, \dots, \ell - 1. \end{aligned}$$

Weiter sei, nicht unbedingt notwendig, aber wünschenswert

$$\begin{aligned} \tau_{i+1} &\leq \tau_i + 1, \\ \chi_{i+1} &\leq \chi_i + 1. \end{aligned}$$

Definition 6.13. Mit den oben angeführten Definitionen wird

$$\mathcal{S}^{(i)} = \left(t^{(\tau_i)}, x^{(\chi_i)}, A^{(\tau_i, \chi_i)} \right)$$

ein Grobgitter bezüglich (\bar{t}, \bar{x}, A) genannt.

Eine Folge

$$\mathcal{S}^{(\ell)} \prec \mathcal{S}^{(\ell-1)} \prec \dots \prec \mathcal{S}^{(0)} = \left(t^{(0)}, x^{(0)}, A \right) = (\bar{t}, \bar{x}, A)$$

wird als Grobgitterhierarchie von A bezeichnet mit der Relation

$$\begin{aligned} \mathcal{S}^{(i)} \prec \mathcal{S}^{(j)} &:\Leftrightarrow \left\{ t^{(\tau_i)} \right\} \subseteq \left\{ t^{(\tau_j)} \right\} \text{ und } \left\{ x^{(\chi_i)} \right\} \subseteq \left\{ x^{(\chi_j)} \right\} \text{ und} \\ &\quad \left| \left\{ t^{(\tau_i)} \right\} \right| + \left| \left\{ x^{(\chi_i)} \right\} \right| < \left| \left\{ t^{(\tau_j)} \right\} \right| + \left| \left\{ x^{(\chi_j)} \right\} \right| \\ &\Leftrightarrow i > j. \end{aligned}$$

Bemerkung 6.14. Die Relation \prec aus Definition 6.13 ist mit den Festlegungen aus Punkt 3 und Punkt 4 transitiv und antisymmetrisch.

Bemerkung 6.15. Die Folgen τ und χ definieren die Reihe der Ausdünnung vom größten Gitter $\mathcal{S}^{(\ell)} = (t^{(l_t)}, x^{(l_x)}, A^{(l_t, l_x)})$ hin zu den Originaldaten.

Bemerkung 6.16. Bei Anwendung einer anderen Folge P als $P_i = 2^i$ muss die Relation \prec aus Definition 6.13 möglicherweise geändert werden, damit weiter eine Grobgitterhierarchie nach obiger Definition aufbaubar ist.

Nomenklatur in den einzelnen Dünngittern

Durch die Verwendung unterschiedlicher Diskretisierungsgitter, ergibt sich für jede Stufe eine eigene Singulärwertzerlegung, eine eigene Zielfunktion und jeweils eine eigene Niedrigrangapproximation. Dazu seien folgende Notationen getroffen.

1. Die Singulärwertzerlegung der Matrix $A^{(\tau_i, \chi_i)}$ in Stufe i :

$$A^{(\tau_i, \chi_i)} = U^{(i)} \Sigma^{(i)} V^{(i)T}. \quad (6.26)$$

2. Die Zielfunktion aus (6.11) für die Diskretisierung aus Stufe i :

$$F^{(i)} : \mathbb{R}^{s \times z} \rightarrow \mathbb{R}. \quad (6.27)$$

3. Die Folge von Iterierten bei der Optimierung von $F^{(i)}(T)$:

$$T^{(i,j)} \in \mathbb{R}^{s \times z}, \quad j = 0, 1, 2, \dots, N_i. \quad (6.28)$$

Dabei ist $T^{(i,0)}$ die Startmatrix für den Minimierungsalgorithmus in der i -ten Stufe, N_i die Anzahl der Iterationen und $\hat{T}^{(i)} = T^{(i,N_i)}$ die letzte Iterierte.

4. Die sich ergebende Niedrigrangapproximation sowie die Transformation $\hat{T}^{(i)}$ in Stufe i :

$$\begin{aligned} (\hat{C}^{(i)}, \hat{A}^{(i)}) &\in \mathbb{R}^{\lfloor \frac{k}{2^i} \rfloor \times s} \times \mathbb{R}^{s \times \lfloor \frac{n}{2^i} \rfloor}, \quad \hat{T}^{(i)} \in \mathbb{R}^{s \times z}, \\ \hat{C}^{(i)} &= U^{(i)} \Sigma_{:,1:z}^{(i)} (\hat{T}^{(i)})^+, \quad \hat{A}^{(i)} = \hat{T}^{(i)} (V_{:,1:z}^{(i)})^T. \end{aligned}$$

Prolongation und Restriktion zwischen zwei Dünngittern

Die Übergänge von einem Grobgitter zu einem benachbarten werden durch Prolongationen und Restriktionen vollzogen.

Definition 6.17. Eine Abbildung vom feinen ins gröbere Gitter

$$\varrho : \mathcal{S}^{(i)} \rightarrow \mathcal{S}^{(i+1)}$$

wird als Restriktion bezeichnet, eine vom groben ins feinere Gitter

$$\rho : \mathcal{S}^{(i+1)} \rightarrow \mathcal{S}^{(i)}$$

als Prolongation.

Bemerkung 6.18. Die Hintereinanderausführung von zunächst Prolongation und danach Restriktion ergibt die Identitätsabbildung, die Umkehrung gilt in der Regel nicht

$$\varrho \circ \rho = \text{id}, \quad \rho \circ \varrho \neq \text{id}.$$

Der Prolongationsoperator

Der Operator ρ realisiert den Übergang von einem groben Gitter $\mathcal{S}^{(i)}$ in das nächstfeinere Gitter $\mathcal{S}^{(i-1)}$. Das Entscheidende ist dabei die Prolongation der Niedrigrangapproximation

$$(C^{(i)}, \hat{A}^{(i)}) \rightarrow (C^{(i-1)}, \hat{A}^{(i-1)}),$$

also der Transformation

$$\hat{T}^{(i)} \rightarrow T^{(i-1,0)}.$$

Als Möglichkeiten bieten sich mit den Abbildungen $\phi_i : \mathbb{R}^{s \times z} \rightarrow \mathbb{R}$, $i = 1, 2, 3$,

$$\begin{aligned} T^{(i-1,0)} &= \arg \min_T \phi_1(T), & \phi_1(T) &= \left\| \hat{A}^{(i)} - T (V_{1:2^{(x_i-x_{i-1})} : \lfloor \frac{n}{2^i} \rfloor, 1:z}^{(i-1)})^T \right\|_F^2, \\ T^{(i-1,0)} &= \arg \min_T \phi_2(T), & \phi_2(T) &= \left\| \hat{C}^{(i)} - U_{1:2^{(\tau_i-\tau_{i-1})} : \lfloor \frac{k}{2^i} \rfloor, :}^{(i-1)} \Sigma_{:,1:z}^{(i-1)} T^+ \right\|_F^2 \end{aligned}$$

oder

$$T^{(i-1,0)} = \arg \min_T \phi_3(T), \quad \phi_3(T) = \phi_1(T) + \phi_2(T)$$

an.

Bemerkung 6.19. In der Praxis ist es sinnvoll die Anpassung nach Art der Prolongation zwischen den Grobgittern $\mathcal{S}^{(i)} \rightarrow \mathcal{S}^{(i-1)}$ auszuwählen

ϕ_1 bei $\tau_i - \tau_{i-1} = 0$, $\chi_i - \chi_{i-1} \geq 1$ (Ausdünnung in x -Richtung),

ϕ_2 bei $\tau_i - \tau_{i-1} \geq 1$, $\chi_i - \chi_{i-1} = 0$ (Ausdünnung in t -Richtung),

ϕ_3 bei $\tau_i - \tau_{i-1} \geq 1$, $\chi_i - \chi_{i-1} \geq 1$ (Ausdünnung in t - und in x -Richtung).

Somit wird diejenige Skala angepasst, in welcher die Prolongation erfolgt. Weiter werden die zugehörigen Prolongationsoperatoren dementsprechend mit ρ_i , $i = 1, 2, 3$, bezeichnet.

Durch den Strukturverlust der Datenmatrix verändern sich die Strukturen der Matrizen der Singulärwertzerlegung (für den Fall $\sigma_j^{(i)} > \sigma_{j-1}^{(i)}$, $i = 1, \dots, \ell$, $j = 1, \dots, z$) ebenfalls wenig. Beim Anpassen der Niedrigrangapproximation der groben Stufe an die nächstfeinere bleibt damit auch die Struktur der Transformation T ebenfalls annähernd erhalten. Für einen Vergleich zweier Matrizen $\hat{T}^{(i)}$ und $\hat{T}^{(i-1)}$ sind jedoch folgende Bemerkungen zu beachten.

Bemerkung 6.20. Für zwei unterschiedliche Dünngitterdiskretisierungen in x entsteht durch die Normierung der rechtsseitigen Singulärvektoren und der Restriktion der Normierung der Zeilen von \hat{A} bezüglich der Maximumsnorm aus f_1 eine andere Skalierung für die Zeilen von T . Für einen sinnvollen Vergleich sollte demzufolge zusätzlich der Faktor $2^{-0.5(\chi_i - \chi_{i-1})}$ einfließen, so dass

$$\hat{T}^{(i)} \approx \frac{1}{2^{0.5(\chi_i - \chi_{i-1})}} \hat{T}^{(i-1)}.$$

Bemerkung 6.21. Für einen Vergleich zweier Transformationen $\hat{T}^{(i)}$ wie in Bemerkung 6.20 beschrieben, sind die Vorzeichen der Singulärvektoren zu beachten. Für den (bei Messwerten) Normalfall $\sigma_j^{(i)} > \sigma_{j+1}^{(i)}$, $i = 0, \dots, \ell$, $j = 1, \dots, z - 1$, sei dies, die Allgemeinheit nicht einschränkend, durch folgende Vorzeichenwahl der Singulärvektoren geregelt

$$\left\| V_{:,j}^{(i)} - V_{1:2^{(\chi_i - \chi_{i-1})}:\lfloor \frac{n}{2^{\chi_i}} \rfloor, j}^{(i-1)} \right\| \leq \left\| V_{:,j}^{(i)} + V_{1:2^{(\chi_i - \chi_{i-1})}:\lfloor \frac{n}{2^{\chi_i}} \rfloor, j}^{(i-1)} \right\|, \quad j = 1, \dots, z.$$

Die algorithmische Struktur

Algorithmisch ist der Ablauf wie folgt organisiert:

1. Restriktion der Daten in ℓ Stufen und auf manuell vorgegebene Weise durch τ und χ .
2. Auf dem größten Gitter werden Optimierungen mit m Zufallsstartmatrizen ausgeführt, welche in der Regel in lokale Minimalstellen führen. Unter diesen wird diejenige mit dem geringsten Funktionswert ausgesucht.
3. Mittels einer der Prolongationsoperatoren ρ_i , $i \in \{1, 2, 3\}$, wird diese zur Startiterierten $T^{(\ell-1,0)}$ im Grobgitter $\mathcal{S}^{(\ell-1)}$ prolongiert.
4. Ausgehend davon wird eine Minimalstelle $\hat{T}^{(\ell-1)}$ bestimmt und diese auf das nächstfeinere Gitter abgebildet. Dieses Vorgehen wird hierarchisch bis zur Wurzelstufe $\mathcal{S}^{(0)}$ fortgesetzt.
5. Letztendlich ergibt sich die Niedrigrangapproximation aus (\hat{C}, \hat{A}) in der Grunddiskretisierung.

Die Abbildung D.1 zeigt diesbezüglich ein Strukturgramm.

Eine Untersuchung der Funktionsweise der Multi-Level-Strategie, insbesondere der Folgen $T^{(i,j)}$, $i = \ell, \dots, 0$, $j = 0, \dots, N_i$, sowie den Minimalstellen $\hat{T}^{(i)}$, $i = \ell, \dots, 0$, erfolgt in Anhang B.5.1 anhand eines Beispiels.

Zusammenfassung 6.22. Die Multi-Level-Strategie dient nur zur Reduktion des Rechenaufwandes. Die Dimension der zu optimierenden Matrix T und damit die Anzahl der Variablen bleibt unverändert. Eine Verbesserung der Stabilität wird erzielt, da die Optimierung in gleicher Zeit öfter mit zufällig gewählten Startvektoren durchlaufen werden kann.

6.5 Strategie der sukzessiven Faktorisierung

In Abschnitt 5.3 ist eine Vorgehensweise zur Bestimmung einer lokalen Rang-1 Approximation einer Matrix beschrieben. Es sind einige Bedenken bezüglich der theoretischen Grundlage für die Ideen von Malinowski und den Algorithmus BTEM genannt, da eine spezielle Datenstruktur für den Erfolg dieser Methode erforderlich ist.

Daten aus der FT-IR Spektroskopie erfüllen unter Umständen diese Voraussetzungen, was auch der Grund für den Erfolg von BTEM bei Daten dieser Spektroskopiertechnik ist.

Ausgehend davon liegt der Ansatz nahe, ähnliches mit den neuen Ideen von PCD zu verknüpfen. Im vorherigen Abschnitt 6.4 erfolgte eine Dimensionsreduktion des Problems in der Auswertung der Zielfunktionen, nicht aber in den Variablen. In diesem Abschnitt wird eine Möglichkeit untersucht, eine Faktorisierung Stück für Stück zu bestimmen. Es werden nacheinander einzeln die Faktorenpaare isoliert und abgespalten. Somit wird eine additive Zerlegung der Matrix A in Einzelschritten

$$A = \underbrace{\hat{C}_{:,1}\hat{A}_{1,:}}_{\text{Schritt 1}} + \underbrace{\hat{C}_{:,2}\hat{A}_{2,:}}_{\text{Schritt 2}} + \cdots + \underbrace{\hat{C}_{:,s}\hat{A}_{s,:}}_{\text{Schritt } s} + E, \quad (6.29)$$

konstruiert. Dieser Ansatz bedeutet demnach eine Variablenreduktion.

Herauslösen eines Faktorenpaars und die lokale Niedrigrangapproximation einer Teilmatrix

Zentraler Punkt der Idee der sukzessiven Matrixfaktorisierung ist die lokale Niedrigrangapproximation durch ein dyadisches Produkt und die Ausweitung dieser lokalen Niedrigrangapproximation auf die gesamte Diskretisierung. Es ergibt sich eine zu Faktorisierungsaufgabe 2.5 abgewandelte Form.

Faktorisierungsaufgabe 6.23. Zu einer gegebenen Matrix $A \in \mathbb{R}^{k \times n}$, $A_{i,j} \geq -\varepsilon$ mit $0 \leq \varepsilon \ll \max_{i,j} A_{i,j}$ und einer Menge $I \subset \{1, \dots, n\}$ werden zwei Vektoren $\hat{C} \in \mathbb{R}_+^k$ und $\hat{A} \in \mathbb{R}_+^{1 \times n}$ mit

$$\|A_{:,I} - \hat{C}\hat{A}_{1,I}\|_F^2 + \sum_{i=1}^p \gamma_i^2 f_i(\hat{C}, \hat{A}) \rightarrow \min$$

gesucht.

Sei die Indexmenge I nun so gewählt, dass $A_{:,I}$ eine Linearkombination aus genau einem Teilfaktor von \hat{A} ist. Dies sei beispielsweise $\hat{A}_{1,:}$. Das heißt es gelte $\|\hat{A}_{1,I}\| > 0$, $\|\hat{A}_{2:s,I}\| = 0$ und

$$A_{:,I} = \hat{C}_{:,1}\hat{A}_{1,I}.$$

Laut (4.5) ergibt sich das Herauslösen dieses Faktorenpaars durch

$$\hat{C}_{:,1} = U\Sigma_{:,1:s}S_{:,1}, \quad \hat{A}_{1,:} = T_{1,:}V_{:,1:s}^T \quad (6.30)$$

und es gilt somit

$$A_{:,I} = \underbrace{U\Sigma_{:,1:s}S_{:,1}}_{\hat{C}_{:,1}} \underbrace{T_{1,:}V_{:,1:s}^T}_{\hat{A}_{1,I}}. \quad (6.31)$$

Zusammenhang zwischen den Linearkombinationen der beiden Faktoren

Unter den Voraussetzungen $\|\hat{A}_{1,I}\| > 0$, $\|\hat{A}_{2:s,I}\| = 0$ ergibt sich (bei ähnlicher Anwendung wie in Gleichung (5.14) durch den BTEM-Algorithmus) zu einem gegebenen Teilfaktor $\hat{A}_{1,:} \in \mathbb{R}^{1 \times n}$ mit $\|\hat{A}_{1,I}\| > 0$ die zugehörige Spalte $\hat{C}_{:,1} \in \mathbb{R}^k$ direkt über

$$\begin{aligned}\hat{C}_{:,1} &= A_{:,I}(\hat{A}_{1,I})^+ \\ &= \frac{1}{\|\hat{A}_{1,I}\|_2^2} A_{:,I} \hat{A}_{1,I}^T.\end{aligned}\tag{6.32}$$

Unverändert wird (für den idealisierten Fall der exakten Matrixfaktorisierung $A = XY$) der Teilfaktor $\hat{A}_{1,:}$ als Linearkombination aus den ersten s rechtsseitigen Singulärvektoren

$$\hat{A}_{1,:} = T_{1,:} V_{:,1:s}^T$$

bestimmt. Für S aus (6.30) folgt nach obigen Überlegungen

$$\begin{aligned}S_{:,1} &= (\Sigma_{:,1:s})^+ U^T \hat{C}_{:,1} = (\Sigma_{:,1:s})^+ U^T \frac{1}{\|\hat{A}_{1,I}\|_2^2} A_{:,I} \hat{A}_{1,I}^T \\ &= \frac{1}{\|T_{1,:} V_{I,1:s}^T\|_2^2} (V_{I,1:s})^T V_{I,1:s} T_{1,:}^T,\end{aligned}$$

was sich deutlich von (5.14) aus dem in Abschnitt 5.3 vorgestellten Algorithmus unterscheidet. Für gestörte Daten wird $\hat{A}_{1,:} = T_{1,:} V_{:,1:s}^T$ und $\hat{C}_{:,1} = U \Sigma_{:,1:s} S_{:,1}$ mit

$$S_{:,1} = \|T_{1,:} V_{I,1:s}^T\|_2^{-2} (V_{I,1:s})^T V_{I,1:s} T_{1,:}^T\tag{6.33}$$

angewendet.

Bemerkung 6.24. Die Forderung $\|\hat{A}_{1,I}\| > 0$, $\|\hat{A}_{2:s,I}\| = 0$ lässt sich vor Beginn der Rechnung nicht belegen. Ein notwendiges aber nicht hinreichendes Kriterium dafür ist

$$\text{rank}(A_{:,I}) = 1$$

oder für gestörte Daten

$$\text{rank}_\varepsilon(A_{:,I}) = 1, \quad 0 < \varepsilon \ll \sigma_1(A_{:,I}).$$

Störungen der lokalen Rang-1-Niedrigrangapproximation

Dass die lokale Niedrigrangapproximation aus Faktorisierungsaufgabe 6.23 für den Fall einer Menge I mit $\exists i_0, i_1, i_0 \neq i_1 : \|\hat{A}_{i_0,I}\|, \|\hat{A}_{i_1,I}\| > 0$ als Faktorisierung keine nützlichen Ergebnisse liefert, ist klar.

Hinsichtlich der Anwendung des Zusammenhangs aus (6.32) für den Einfluss von Störungen durch die Teilfaktoren $\hat{A}_{2:s,:}$ auf der Indexmenge I gibt folgender Satz eine Aussage.

Satz 6.25. Sei I ein Indexbereich mit $\|\hat{A}_{1,I}\| > 0$. Weiter seien $\hat{A}_{1,I} = T_{1,:} V_{I,1:s}^T$ exakt bestimmt und die Gewichtung des zugehörigen Faktors

$$\tilde{C}_{:,1} = \|\hat{A}_{1,I}\|_2^{-2} A_{:,I} \hat{A}_{1,I}^T$$

wie in (6.32) berechnet. Auf dem Indexbereich I liege eine Störung

$$A_{:,I} = \hat{C}_{:,1} \hat{A}_{1,I} + \underbrace{\sum_{i=2}^s \hat{C}_{:,i} \hat{A}_{i,I}}_E$$

vor, wobei \hat{C} und \hat{A} die exakten Faktoren seien.

Dann gilt für die Abweichung der approximierten Spalte $\tilde{C}_{:,1}$ zur exakten $\hat{C}_{:,1}$

$$\begin{aligned} \max_{i=2,\dots,s} \hat{A}_{i,I}^T \hat{A}_{1,I} \|\hat{C}_{:,i}\| &\leq \|\hat{A}_{1,I}\|_2^2 \|\tilde{C}_{:,1} - \hat{C}_{:,1}\| \leq \sum_{i=2}^s \hat{A}_{i,I}^T \hat{A}_{1,I} \|\hat{C}_{:,i}\|, \\ \|\hat{C}_{:,1} - \tilde{C}_{:,1}\|_2 &\leq \frac{\|E\|_2}{\|\hat{A}_{1,I}\|_2}. \end{aligned} \quad (6.34)$$

Beweis: Da $\|\hat{A}_{1,I}\| > 0$ ist, folgt

$$\begin{aligned} \tilde{C}_{:,1} &= \left(\sum_{i=1}^s \hat{C}_{:,i} \hat{A}_{i,I} \right) \hat{A}_{1,I}^+, \\ \tilde{C}_{:,1} - \hat{C}_{:,1} &= \left(\sum_{i=2}^s \hat{C}_{:,i} \hat{A}_{i,I} \right) \hat{A}_{1,I}^+, \\ \|\tilde{C}_{:,1} - \hat{C}_{:,1}\|_2 &\leq \|E\|_2 \|\hat{A}_{1,I}^+\|_2 \leq \frac{\|E\|_2}{\|\hat{A}_{1,I}\|_2}. \end{aligned} \quad (6.35)$$

Die Ungleichungen aus (6.34) ergeben sich über (6.35) durch die Nichtnegativität von \hat{C} und \hat{A} (sowie somit auch $\hat{A}_{1,I}^+$). \square

Bemerkung 6.26. Im Vergleich zu den Abschätzungen aus (5.21) bezüglich des Algorithmus BTEM sind die Abschätzungen aus (6.34) für nichtnegative Faktoren \hat{C}, \hat{A} deutlich schärfer. Insbesondere ist $\tilde{C} = \hat{C}$, sofern $\hat{A}_{i,I}^T \hat{A}_{1,I} = 0$ für alle $i = 2, \dots, s$ gilt.

Der Algorithmus der sukzessiven Niedrigrangapproximation

Insgesamt ergibt sich damit als Ablauf eine Iteration über die Anzahl der Komponenten. Es wird jeweils zu einem geeigneten Indexbereich I über Faktorisierungsaufgabe 6.23 und (6.32) ein Faktorenpaar $(\hat{C}_{:,i}, \hat{A}_{i,:})$ bestimmt, anschließend abgespalten und die Iteration über

$$A \leftarrow A - \hat{C}_{:,i} \hat{A}_{i,:}$$

fortgesetzt. Schlussendlich bleiben Störungen übrig. Abbildung D.2 im Anhang zeigt ein entsprechendes Ablaufdiagramm.

6.5.1 Kritische Bewertung des Ansatzes der sukzessiven Zerlegung

Trotz allen praktischen Erfolgs bei der Anwendung von BTEM auf Daten aus der FT-IR Spektroskopie und den oben angeführten sinnvollen Evolutionen sei ausdrücklich darauf hingewiesen, dass starke Überlagerungen in den Zeilen von \hat{A} auf dem Indexausschnitt I zu erheblichen Fehlern in der Niedrigrangapproximation führen können. Es ist daher von der Anwendung dieser Methode abzuraten und die Aufgabe als solche zu betrachten, die sie auch ist: Die Auflösung eines komplexen gekoppelten Systems.

Die Subtraktion einzelner Faktorenpaare bei gestörten Daten hat numerische Fehler zur Folge, welche die weitere Niedrigrangapproximation der verbleibenden Matrix $A - \hat{C}_{:,i_0} \hat{A}_{i_0,:}$ stören. Umgehen lässt sich dieses Problem mit der Verfahrensweise von BTEM, die isolierten Faktorenpaare nur zu bestimmen und die Ermittlung weiterer Komponenten ohne Subtraktion vorheriger vorzunehmen. Das heißt für alle $i = 1, \dots, s$ werden die Faktorenpaare ausgehend von den Originaldaten A berechnet. Speziell der Umgang mit Faktorenpaaren, welche einen verhältnismäßig geringen Beitrag zur Matrix A liefern, wird durch den oben beschriebenen Vorgang erschwert.

6.6 Gewichtung der Regularisierungsfunktionen

Reduziert sich die Menge aller Faktorisierungen nicht auf eine Äquivalenzklasse, so erfolgt der Übergang zur regularisierten nichtnegativen Niedrigrangapproximation. Dabei bestimmen die Gewichte $\gamma_i \in \mathbb{R}$, $i = 1, \dots, p$, entscheidend die resultierenden Faktoren, da verschiedene Wahlen von γ im Normalfall auch auf Faktorisierungen aus unterschiedlichen Äquivalenzklassen führen. Der Einfluss dieser Gewichte wird in diesem Abschnitt diskutiert.

Die Zielfunktion

Wie in Abschnitt 6.1 erläutert, setzt sich die Zielfunktion aus (6.12) als

$$F(T) = \frac{1}{2} \left(\bar{f}_0^2(T) + \sum_{i=1}^p \gamma_i^2 \bar{f}_i^2(T) \right),$$

zusammen. Vereinfachend seien die Matrix T zeilenweise hintereinander aufgelistet, in den Vektor x überführt und F dementsprechend über

$$\begin{aligned} \phi : \mathbb{R}^{sz} &\rightarrow \mathbb{R}^{s \times z}, & (\phi(x))_{i,j} &= x_{(i-1)z+j}, & i &= 1, \dots, s, & j &= 1, \dots, z, \\ g_i : \mathbb{R}^q &\rightarrow \mathbb{R}, & g_i(x) &= \bar{f}_i(\phi(x)), & i &= 0, \dots, p, \end{aligned}$$

angepasst durch

$$G : \mathbb{R}^q \times \mathbb{R}^p \rightarrow \mathbb{R}, \quad G(x; \gamma) = \frac{1}{2} \left(g_0^2(x) + \sum_{i=1}^p \gamma_i^2 g_i^2(x) \right).$$

Dabei ist $q = sz$ die Anzahl der Variablen.

Es stellt sich die Frage nach einer geeigneten Gewichtung und Auswahl der Einzelfunktionen $g_i(x)$. Ziel ist einerseits $G(x; \gamma)$ zu einer festen Parameterwahl γ und in Abhängigkeit von x zu minimieren und andererseits, dass die Minimalstelle $x = \hat{x} \in \mathbb{R}^{sz}$ eine geeignete Niedrigrangapproximation (\hat{C}, \hat{A}) der Matrix A ergibt. Dabei hängt die Minimalstelle \hat{x} direkt von der Wahl der Gewichtung γ ab

$$\hat{x} = \hat{x}(\gamma) = \hat{x}_\gamma.$$

Der Einfluss der Regularisierungsparameter

Seien nun \hat{T} eine geeignete Transformation für \hat{C} und \hat{A} , umgeformt zu \hat{x} , und γ eine fest gewählte Gewichtung. Demzufolge sollte diese Gewichtung so gewählt sein, dass \hat{x} eine globale Minimalstelle der zweimal stetig differenzierbaren Funktion $G(x; \gamma)$ unter Fixierung von γ ist. Das heißt notwendigerweise, dass

$$\|\nabla_x G(\hat{x}; \gamma)\| = 0 \tag{6.36}$$

ist sowie hinreichend dazu, dass

$$H_G(\hat{x}; \gamma) \quad \text{positiv definit} \tag{6.37}$$

ist mit

$$\nabla_x G \in \mathbb{R}^q, \quad \nabla_x G(x; \gamma) = \left(\left(\frac{\partial g_0}{\partial x_i} g_0 + \sum_{l=1}^p \gamma_l^2 \frac{\partial g_l}{\partial x_i} g_l \right) \right)_{i=1, \dots, q},$$

und H_G der Hessematrix der zweiten partiellen Ableitungen der Funktion G nach x

$$H_G \in \mathbb{R}^{q \times q}, \quad H_G(x; \gamma) = \left(\frac{\partial^2 G}{\partial x_i \partial x_j}(x; \gamma) \right)_{i,j=1,\dots,q}.$$

Weiter wäre von \hat{x} wünschenswert, dass neben den notwendigen und hinreichenden Forderungen aus (6.36) und (6.37) zusätzlich

$$\|\nabla_\gamma \nabla_x G(\hat{x}; \gamma)\| = 0 \quad (6.38)$$

gilt mit $\nabla_\gamma \nabla_x G(x; \gamma) \in \mathbb{R}^{q \times p}$,

$$(\nabla_\gamma \nabla_x G(x; \gamma))_{i,j} = \frac{\partial^2 G}{\partial x_j \partial \gamma_i}(x, \gamma) = \left(2\gamma_j \frac{\partial g_j}{\partial x_i} g_j \right)(x; \gamma), \quad i = 1, \dots, q, \quad j = 1, \dots, p.$$

Die Problematik der Parameterwahl

Ziel ist es, dass bei einer Veränderung des Parametersatzes γ der Punkt \hat{x} auch weiterhin die notwendige Bedingung eines verschwindenden Gradienten für eine lokale Minimalstelle erfüllt. Es ergibt sich somit die Forderung aus (6.38). Für die Praxis ist das eine nicht sinnvolle Forderung, was die Korollare 6.27 und 6.28 zeigen.

Korollar 6.27. *Die Bedingung aus (6.38) ist mit $\min(\gamma) > 0$ genau dann erfüllt, wenn gilt*

$$g_i(\hat{x}) \|\nabla_x g_i(\hat{x}_\gamma)\| = 0, \quad \forall i = 1, \dots, p. \quad (6.39)$$

Die Forderung aus (6.39) ist für lineare Funktionen

$$g_i(x) = (a^{(i)})^T x$$

mit $a^{(i)} \in \mathbb{R}^q$ und für $i = i_0, \dots, i_1$ unter Umständen nur bedingt sinnvoll, da daraus die Forderung $x \in \ker((a^{(i_0)}, \dots, a^{(i_1)})^T)$ folgt. Für die Aufspaltung der Regularisierungsfunktion f_6 aus (6.9) in die einzelnen Quadratkomponenten zu $\bar{f}_{i_0}, \dots, \bar{f}_{i_1}$, ist dies nur erfüllt, falls $x = 0$ gilt, da die zugehörige Matrix $(a^{(i_0)}, \dots, a^{(i_1)})^T$ keinen echten Kern besitzt. Dies gilt in der Regel ebenfalls für f_4 und f_5 aus (6.8) sowie f_7 . Dazu ist folgendes Korollar formuliert.

Korollar 6.28. *Die Forderung aus (6.38) beziehungsweise (6.39) ist für die aus f_4 , f_5 , f_6 und f_7 erzeugten Teilfunktionen g_i nicht sinnvoll.*

Die Aufspaltung von f_4 aus (6.8) in Quadratkomponenten, das Einsetzen der Transformation $\hat{A} = TV_{:,1:z}^T$ aus (6.2) und die Umsetzung in g_i , $i = i_0^{(4)}, \dots, i_1^{(4)}$, mit entsprechend gewählten $i_0^{(4)}, i_1^{(4)}$ ergibt die linearen Funktionen

$$g_{i_0^{(4)} + (i-1)(n-2) + j}(x) = \sqrt{b_j^{(4)}} \sum_{l=1}^n \sum_{\nu=1}^z B_{j,l}^{(4)} V_{l,\nu} x_{(i-1)z + \nu}, \quad i = 1, \dots, s, \quad j = 1, \dots, n-2.$$

Damit gilt die Bedingung aus (6.39) für $i = i_0^{(4)}, \dots, i_1^{(4)}$ lediglich, falls

$$g_{i_0^{(4)} + (i-1)(n-2) + j}(x) = 0$$

ist. Dies gilt in der Regel nur für alle $i = 1, \dots, s$, $j = 1, \dots, n-2$, falls $x = 0$ ist, da

$$\ker(V_{:,1:z}) = \{0\}$$

sowie im Allgemeinen gilt

$$\text{im}(V_{:,1:z}) \not\subseteq \ker(B^{(4)}).$$

Für f_5 gilt Analoges, für f_7 ebenfalls, nur unter Einbeziehung der linksseitigen Singulärvektoren sowie den Singulärwerten. Für f_6 entfällt die Einschränkung „in der Regel“. \square

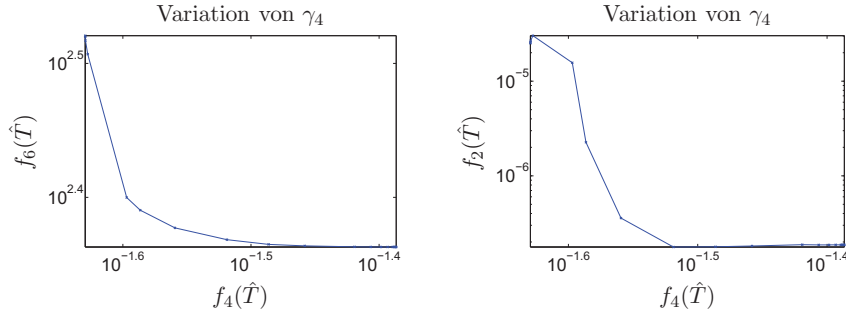


Abbildung 6.1: Bestimmt wurden Faktorisierungen des Datensatzes bezüglich der Fallstudie aus Anhang B.4. Variiert ist die Gewichtung $\gamma_4 \in [0.0025, 25]$ aus (6.11) mit den festen Gewichten $\gamma_1 = 4$, $\gamma_2 = 0.4$, $\gamma_3 = 4$, $\gamma_6 = 0.02$, $\gamma_7 = 0.01$ sowie $\gamma_i = 0$, $i = 5, 8, 9, 10, 11, 12, 13$. In beiden Darstellungen sind die L-Kurven jeweils gut zu erkennen. Im Sinne von dünn besetzten Zeilenvektoren in \hat{A} sind dementsprechend die Werte bezüglich kleineren Residuen für f_6 interessanter.

Das Entscheidungskriterium der L-Kurve

Eine Möglichkeit, die Wahl der Parameter a-posteriori zu bewerten, ist eine sogenannte L-Kurve, [7, 33–35]. Ziel ist es dabei, für eine Quadratmitteloptimierung

$$\min_{x \in \mathbb{R}^q} \frac{1}{2} \left(g_0^2 + \sum_{i=1}^p \gamma_i^2 g_i^2(x) \right)$$

verschiedene Gewichtungssätze γ zu vergleichen, wobei in der Regel nur ein Parameter γ_l variiert wird. Dabei liefern die Parameter $\gamma^{(i)}$, $i = 1, \dots, m$, die optimalen Variablen

$$x^{(i)} = x^{(i)}(\gamma^{(i)})$$

mit

$$G(\hat{x}^{(i)}) = \min_{x \in \mathbb{R}^q} G(x; \gamma^{(i)}), \quad i = 1, \dots, m.$$

Wurde nur ein Parameter γ_l variiert und enthält die doppelt-logarithmische Darstellung der ungewichteten Quadratmittel

$$(g_l^2(\hat{x}^{(i)}), g_\ell^2(\hat{x}^{(i)})), \quad i = 1, \dots, m,$$

für zwei Teilfunktionen g_l , g_ℓ einen „Knick“ und beschreibt die Form des Buchstaben „L“, so wird die Parameterwahl von $\gamma_l^{(i_0)}$ und $\gamma_\ell^{(i_0)}$ zugehörig zu dem „Knick“ als geeignete Wahl für einen Kompromiss aus den beiden Teilfunktionen g_l und g_ℓ angesehen, vergleiche dazu die Abbildungen 6.1 und 6.2.

6.7 Programmstruktur und Implementierung

Basierend auf den Untersuchungen dieses Kapitels ist ein Programm zur Berechnung regularisierter nichtnegativer Niedrigrangapproximationen entstanden, welches bereits am Leibniz-Institut für Katalyse e.V. in Rostock sowie der Evonik Oxeno GmbH in Marl zur Reinkomponentenzerlegung spektroskopischer Daten eingesetzt wird.

Dieser Teilabschnitt umfasst eine kurze Erläuterung des Programmaufbaus. Der Programmablauf ist in drei verschiedene Abschnitte unterteilt, wobei diese in verschiedenen Programmiersprachen implementiert sind.

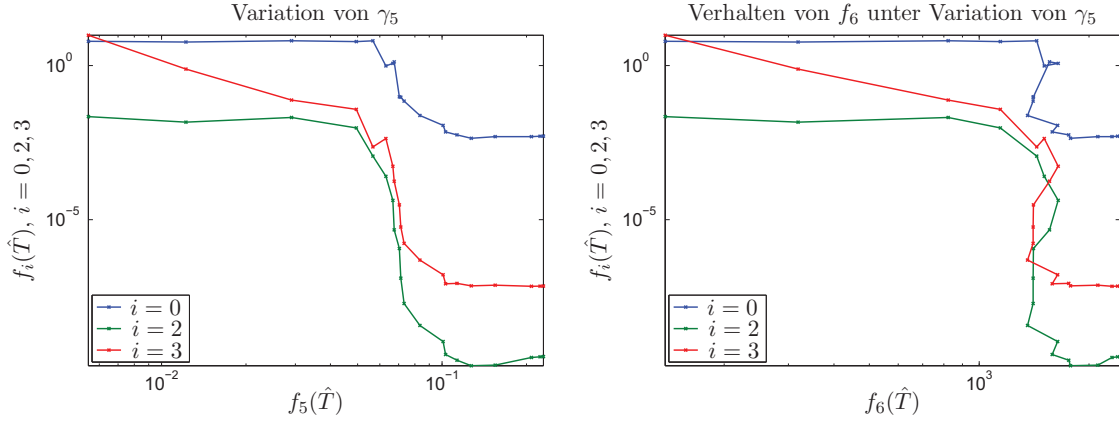


Abbildung 6.2: Bestimmt wurden Faktorisierungen des Datensatzes bezüglich der Fallstudie aus Anhang B.2. Variiert ist die Gewichtung $\gamma_5 \in [0.0015, 150]$ aus (6.11) bei den festen Gewichten $\gamma_1 = 1$, $\gamma_2 = 0.1$, $\gamma_3 = 1$, $\gamma_4 = 0.01$, $\gamma_7 = 10^{-6}$, $\gamma_{12} = 0.0025$ und $\gamma_i = 0$, $i = 5, 8, 9, 10, 11$. Entscheidend ist die linke Darstellung, in der die L-Kurven gut zu erkennen sind. Ergänzend dazu lässt die rechte Darstellung eine Korrelation mit der Entwicklung von $f_6(\hat{T})$ für dieses Beispiel erkennen. Bei der Betrachtung der Faktoren \hat{C} und \hat{A} ist die Parameterwahl des „Knicks“ der L-Kurve ebenfalls im Sinne der Chemie optimal.

Dabei ist in einen MATLAB-Code ein C-Unterprogramm eingebunden, welches FORTRAN Bibliotheken verwendet. Die MATLAB-Datei fungiert als Ausgangspunkt von PCD und der C-Code als Geschwindigkeitsverbesserung der rechenintensiven Optimierung. Die FORTRAN Bibliotheken dienen sowohl als eigentlicher Optimierer als auch zur Bereitstellung mathematischer Hilfsmittel.

Der Gesamttablauf erfolgt aus MATLAB heraus, dazu zählt die Datenverarbeitung und das Einlesen aller notwendigen Parameter (Indexbereiche der möglicherweise angewendeten lokalen Niedrigrangapproximation, s , z , Angaben zur Multi-Level-Strategie usw.). Auch die Multi-Level-Strategie wird aus MATLAB heraus gesteuert. Dazu werden eine Singulärwertzerlegung der betrachteten Matrix generiert und alle notwendigen Matrizen und Vektoren bestimmt. Anschließend erfolgt der Aufruf der C-Routine PURECOMPONENTDECOMPOSITION, welche den Teil der Optimierung umfasst. Die Schnittstelle zwischen dem MATLAB- und dem C-Code bilden dabei Dateien, welche mit ausreichender Genauigkeit (Exponentialdarstellung, 16-stelliger Mantisse) ex- und importiert werden. In dem C-Code werden alle rechen- und damit zeitaufwendigen Operationen ausgeführt, also insbesondere die Auswertung der Zielfunktion, welche Bestandteil der C-Routine ist. In dem C-Code eingebunden sind die FORTRAN Bibliotheken LAPACK und BLAS. Die Optimierung der Zielfunktion wird mittels der in Abschnitt 6.3 erläuterten ACM FORTRAN Bibliothek NL2SOL ausgeführt. Von dem C-Code wird eine Transformation T zurückgegeben, welche in der Regel ein lokales Minimum darstellt. Die Weiterverarbeitung erfolgt in dem MATLAB-Code, welcher schlussendlich mit der Ausgabe von \hat{C} und \hat{A} terminiert. Für eine Darstellung der Verknüpfungen der einzelnen Teilfunktionen ist auf das Diagramm aus Abbildung D.3 verwiesen.

6.8 Kritische Bewertung des Algorithmus

Ansatz der Transformation der Singulärvektoren

Der Ansatz, die Faktorisierungsaufgabe nicht direkt durch eine Berechnung der Faktoren \hat{C} und \hat{A} zu lösen, sondern mit Hilfe einer Transformation der Singulärvektoren wie in (6.2), verringert den Rechenaufwand für große Unterschiede von $\max\{s, z\}$ und $\min\{k, n\}$ enorm.

Für leicht gestörte Daten ist die Einschränkung sich möglicherweise ergebender kleiner negativer Einträge in den Faktoren und die Beschränkung der Freiheitsgrade akzeptabel, da sich die signifikanten Charakteristiken für die Niedrigrangapproximation in den ersten Singulärvektoren wiederfinden.

Ist die Störung E im Vergleich zum Produkt $\hat{C}\hat{A}$ stark, so ist der Ansatz die Faktorisierung auf Basis einer abgeschnittenen Singulärwertzerlegung der gestörten Matrix $A = \hat{C}\hat{A} + E$ zu berechnen, nicht zu empfehlen. Es ist in diesem Fall nicht gesichert, dass es eine Transformation der Singulärvektoren zu Faktoren gibt, welche einerseits nur vergleichsweise geringe negative Einträge besitzen und deren Produkt andererseits eine „akzeptable“ Näherung an A darstellt.

Multi-Level-Methoden und Zufallsstartmatrizen

Das zur Berechnung der regularisierten Niedrigrangapproximation resultierende Optimierungsproblem besitzt für $s \geq 2$ im Normalfall keine eindeutige globale Minimalstelle und viele lokale Minimalstellen. Um die Wahrscheinlichkeit zu minimieren, ein lokales und nicht globales Minimum zu erhalten, wird die Optimierung mehrfach mit zufällig gewählten Startmatrizen $T \in \mathbb{R}^{s \times z}$ durchgeführt.

Dem gesteigerten Rechenaufwand lässt sich wirkungsvoll durch eine Multi-Level-Strategie entgegensteuern. Voraussetzung dafür sind strukturierte Daten, welche sich aus der Diskretisierung des Faktorisierungsproblems 2.2 unter bestimmten Bedingungen ergeben.

Sukzessives Abspalten der Faktoren

Die Berechnung einer Niedrigrangapproximation durch eine sukzessive Faktorisierung mittels lokaler Niedrigrangapproximationen durchzuführen, ist kein geeigneter Ansatz, die Aufgabe 2.5 zufriedenstellend zu lösen. Überlagerungen, sowohl in dem lokal betrachteten Bereich, als speziell auch außerhalb diesen, werden so ignoriert. Ohne analytische Grundlage wird unterstellt, dass das verbleibende Residuum $A - \hat{C}_{:,i}\hat{A}_{i,:}$ eine nichtnegative Niedrigrangapproximation vom Rang $s - 1$ mit physikalisch sinnvollen Faktoren besitzt.

Wahl der Gewichtung

Die Uneindeutigkeit der nichtnegativen Matrixfaktorisierung ist die Motivation für den Übergang zur regularisierten Niedrigrangapproximation. Dabei stellt sich die Frage nach der Wahl der Regularisierungsparameter γ_i , $i = 1, \dots, p$. Da das Minimum der Zielfunktion

$$f_\gamma(x) = G(x, \gamma)$$

direkt von γ abhängt ($\hat{x} = \hat{x}(\gamma)$), nimmt die Parameterwahl eine Schlüsselrolle ein. Nicht mehr als eine Hilfestellung bietet eine zu speziell gewählten Parametersätzen erstellte L-Kurve, welche Hinweise auf eine geeignete Parameterrelation liefert.

Perspektiven

Die Perspektiven der in diesem Kapitel behandelten Regularisierungen sind begrenzt. Das Problem der nichtnegativen Niedrigrangapproximation lässt sich durch Regularisierungsfunktionen beeinflussen. Dieser Einfluss ist mit der Auswahl der Regularisierungen jedoch abhängig von der zugrunde gelegten Modellierung.

Die untersuchte Faktorisierungsaufgabe mit den bekannten Einschränkungen der Nichtnegativität ist und bleibt nicht eindeutig lösbar, lediglich werden mit Hilfe von Regularisierungen aus den möglichen Faktorisierungen bestimmte selektiert.

7 Kinetische Modellanpassung

In den vorherigen Kapiteln wurden multivariate Modelle der Form $A = \hat{C}\hat{A} + E$ und Algorithmen zur Berechnung der Faktoren \hat{C} und \hat{A} untersucht. In diesem Kapitel wird eine solche Faktorisierung zur Anpassung eines kinetischen Modells genutzt.

Ausgangspunkt ist eine nichtnegative Niedrigrangapproximation aus Faktoren \hat{C}, \hat{A} , wobei die Einträge der Ausgangsmatrix A spektroskopische Daten einer chemischen Reaktion sind. Die Einträge des Faktors \hat{C} sind Approximationen an die vorliegenden Einzelkonzentrationen zu den Zeitpunkten $\bar{t}_1, \dots, \bar{t}_k$, also diskrete Auswertungen einer Funktion $c \in \mathcal{C}(\mathbb{R}_+, \mathbb{R}_+^s)$ von s Konzentrationsverläufen an den Stellen $\bar{t}_1, \dots, \bar{t}_k$.

In der Chemie ist die Reaktionskinetik eines Systems aus s Komponenten durch eine zugehörige gewöhnliche Differentialgleichung $\dot{c} = f(t, c)$ beschrieben. Ist die Differentialgleichung bekannt, so sind die Konzentrationsverläufe $c(t)$, $t \in [\bar{t}_1, \bar{t}_k]$, zu dem Anfangswert $c(\bar{t}_1)$ direkt berechenbar. Enthält die Differentialgleichung jedoch einen unbekannten Parameter $\phi \in \mathbb{R}^\nu$,

$$\dot{c} = f(t, c; \phi), \quad (7.1)$$

so lässt sich das Ergebnis der nichtnegativen Matrixfaktorisierung zu dessen Approximation als Lösung einer quadratischen Ausgleichsrechnung nutzen.

Der Parameters ϕ soll so angepasst werden, dass die Lösung des Anfangswertproblems bestmöglich an den Approximationen \hat{C} im Sinne der kleinsten Quadrate liegt. Mit $C^{(\text{dgl})}(\phi)$ als Matrix der Auswertungen der Lösung $c(t)$ des resultierenden Anfangswertproblems und $C_{i,j}^{(\text{abs})}$ den korrekt skalierten Konzentrationsprofilen ist, wie in Abschnitt 6.2 erläutert, die Aufgabe einer Ausgleichsrechnung

$$\sum_{i=2}^k \sum_{j=1}^s (C_{i,j}^{(\text{dgl})}(\phi) - C_{i,j}^{(\text{abs})})^2 \rightarrow \min$$

gestellt.

In Abschnitt 6.2 wurde eine Möglichkeit erläutert, die Bestimmung des Parameters ϕ direkt mit der Berechnung einer Faktorisierung zu koppeln. In diesem Kapitel wird die Parametrisierung eigenständig und im Anschluss an die Berechnung der Faktorisierung vorgenommen. Dabei soll dieses Kapitel eine mögliche Weiterverarbeitung eines der Faktoren erläutern und kann nur als weiterführende Ergänzung angesehen werden. Der Fokus liegt deutlich auf den vorherigen Kapiteln. Für nähere Untersuchungen zur Reaktionskinetik ist auf [72], für Untersuchungen beispielsweise zu den artverwandten Randwertproblemen ist auf [1, 81] verwiesen.

Skalierung

Ein Freiheitsgrad der Niedrigrangapproximation ist die Skalierung. In Abschnitt 6.2.3 wurden verschiedene Möglichkeiten erläutert, diesen zu eliminieren. Im Folgenden sei dieser Schritt mit $\alpha \in \mathbb{R}_+^s$ bereits vollzogen und der Faktor

$$C^{(\text{abs})} = \hat{C} \text{diag}(\alpha)$$

gesetzt.

7.1 Die Anwendung der regularisierten Niedrigrangapproximation

Der extrahierte und skalierte Faktor $C^{(\text{abs})}$ aus der Niedrigrangapproximation von A ist der Ausgangspunkt des zu lösenden Problems. Dieser enthält diskrete Approximationen der Funktion $c(t)$ der Konzentrationsprofile zu dem Zeitgitter \bar{t}_i , $i = 1, \dots, k$,

$$C^{(\text{abs})} \in \mathbb{R}^{k \times s}, \quad C_{i,j}^{(\text{abs})} \approx c_j(\bar{t}_i), \quad i = 1, \dots, k, \quad j = 1, \dots, s.$$

Dem untersuchten System liege ein kinetisches Modell in Form der gewöhnlichen Differentialgleichung

$$\dot{c}(t) = f(t, c; \phi), \quad t \in [\bar{t}_1, \bar{t}_k], \quad (7.2)$$

für den Faktor $c : \mathbb{R} \rightarrow \mathbb{R}^s$ zugrunde, wobei $\phi \in \mathbb{R}^\nu$ der zu bestimmende Parameter ist und $f : \mathbb{R} \times \mathbb{R}^s \times \mathbb{R}^\nu \rightarrow \mathbb{R}^s$ die rechte Seite der Differentialgleichung. Als Anfangswert ist

$$c(\bar{t}_1) = C_{1,:}^{(\text{abs})} \quad (7.3)$$

gegeben.

Zu dem Modell (7.2) ist ein Parameter ϕ derart gesucht, dass die Lösung $c(t)$ zu dem Anfangswert (7.3) die vorliegenden Approximationen des Faktors $(\bar{t}_i, C_{i,:}^{(\text{abs})})$, $i = 2, \dots, k$, im Sinne der kleinsten Quadrate bestmöglich approximiert.

Um die lokale Existenz einer Lösung und deren Eindeutigkeit zu garantieren, wird an das Anfangswertproblem folgende Voraussetzung gestellt.

Voraussetzung 7.1. Für das Anfangswertproblem mit der Differentialgleichung aus (7.2) und dem Anfangswert (7.3) seien zu $[\bar{t}_1, \bar{t}_n]$ eine offene Menge $\Omega \subset \mathbb{R}^s$ mit $C_{1,:}^{(\text{abs})} \in \Omega$, eine Menge $\Phi \subset \mathbb{R}^\nu$ sowie eine Lipschitzkonstante L für die rechte Seite f gegeben, so dass für $f : \Omega \times \Phi \rightarrow \mathbb{R}^s$ gilt

$$\|f(t, c; \phi) - f(t, c'; \phi)\| \leq L\|c - c'\|, \quad \forall t \in [\bar{t}_1, \bar{t}_n], \quad c, c' \in \Omega, \quad \phi \in \Phi.$$

Bemerkung 7.2. Damit existieren nach dem Existenz- und Eindeutigkeitssatz von Picard-Lindelöf, [21], für jede Wahl von $\phi \in \Phi$ ein Wert $\delta > 0$ und lokal eine eindeutig bestimmte Lösung $c(t)$, $t \in [\bar{t}_1, \delta]$, die der Differentialgleichung (7.2) und $c(\bar{t}_1) = C_{1,:}^{(\text{abs})}$ genügt.

7.2 Eine Folge von Anfangswertproblemen

Wie in Abschnitt 6.2 angeführt, ist die nichtlineare Quadratmitteloptimierung der $s(k-1)$ Fehlerfunktionen

$$g_{i+(j-1)s} : \mathbb{R}^\nu \rightarrow \mathbb{R}, \quad g_{i+(j-1)s}(\phi) = C_{j+1,i}^{(\text{dgl})}(\phi) - C_{j+1,i}^{(\text{abs})} = 0, \quad j = 1, \dots, k-1, \quad i = 1, \dots, s, \quad (7.4)$$

eine Möglichkeit der Parametrisierung, wobei $C^{(\text{dgl})}$ die Auswertungen der Lösung $c(t)$ auf dem Gitter \bar{t} enthält. Insgesamt lautet die Zielfunktion $G : \mathbb{R}^\nu \rightarrow \mathbb{R}$ für die Ausgleichsrechnung

$$G(\phi) = \frac{1}{2} \sum_{i=1}^{s(k-1)} g_i^2(\phi).$$

Zur Bestimmung der einzelnen Approximationen $C_{j,i}^{(\text{dgl})}$, $j = 2, \dots, k$, $i = 1, \dots, s$, wurden in Abschnitt 6.2 zwei Möglichkeiten über die Lösung jeweils einer Folge von Anfangswertproblemen aufgezeigt. Dabei boten sich einerseits eine induktiv sowie andererseits eine autonom aufgebaute Folge von Anfangswertproblemen an. Diese Ansätze werden nun ebenfalls zur Lösung der in diesem Kapitel gestellten Aufgabe verwendet.

7.2.1 Eine induktive Folge von Anfangswertproblemen

Mit der Initialisierung $C_{1,:}^{(\text{dgl})} := C_{1,:}^{(\text{abs})}$ wird für $i = 1, \dots, k-1$ eine induktive Folge von Anfangswertproblemen wie folgt definiert

$$\begin{aligned} \dot{c}(t) &= f(t, c; \phi), & t \in [\bar{t}_i, \bar{t}_{i+1}], \\ c(\bar{t}_i) &= C_{i,:}^{(\text{dgl})}. \end{aligned} \quad (7.5)$$

Dabei wird jeweils die Auswertung der Lösung $c(t)$ des Anfangswertproblems für $[\bar{t}_i, \bar{t}_{i+1}]$ am Intervallende, also $c(\bar{t}_{i+1})$, als Anfangswert des folgenden Anfangswertproblems (für $[\bar{t}_{i+1}, \bar{t}_{i+2}]$) gewählt

$$C_{i+1,:}^{(\text{dgl})} := c(\bar{t}_{i+1}).$$

Für den Gradienten von G folgt nach dieser Vorgehensweise

$$\nabla G = J_g^T g, \quad J_g \in \mathbb{R}^{((k-1)s) \times \nu}, \quad (J_g)_{i,\ell} = \frac{\partial g_i}{\partial \phi_\ell}, \quad i = 1, \dots, (k-1)s, \quad \ell = 1, \dots, \nu,$$

wobei für die Jacobimatrix J_g von g mit der Notation $g_{:(j-1)s} = (g_{1+(j-1)s}, \dots, g_{js})^T$ gilt

$$\begin{aligned} \frac{\partial g_{:(j-1)s}}{\partial \phi_\ell} &= \frac{\partial}{\partial \phi_\ell} \left((C_{1,:}^{(\text{dgl})})^T + \int_{\bar{t}_1}^{\bar{t}_{j+1}} f(\tau, c(\tau); \phi) \, d\tau - (C_{j+1,:}^{(\text{abs})})^T \right), & j = 1, \dots, k-1, \\ &= \frac{\partial c}{\partial \phi_\ell}(\bar{t}_{j+1}). \end{aligned} \quad (7.6)$$

Zu beachten ist die indirekte Abhängigkeit der Lösung $c(t)$ von ϕ

$$\begin{aligned} \frac{\partial c}{\partial \phi_\ell}(\bar{t}_{j+1}) &= \int_{\bar{t}_1}^{\bar{t}_{j+1}} \left(\sum_{\mu=1}^s \frac{\partial f}{\partial c_\mu} \frac{\partial c_\mu}{\partial \phi_\ell} + \frac{\partial f}{\partial \phi_\ell} \right) (\tau, c(\tau); \phi) \, d\tau \\ &= \int_{\bar{t}_1}^{\bar{t}_{j+1}} (f_c c_\phi + f_\phi)(\tau, c(\tau); \phi) \, d\tau \end{aligned}$$

mit den partiellen Ableitungen

$$\begin{aligned} f_c &\in \mathbb{R}^{s \times s}, & (f_c)_{i,j} &= \frac{\partial f_i}{\partial c_j}, \\ c_\phi &\in \mathbb{R}^{s \times \nu}, & (c_\phi)_{i,j} &= \frac{\partial c_i}{\partial \phi_j}, \\ f_\phi &\in \mathbb{R}^{s \times \nu}, & (f_\phi)_{i,j} &= \frac{\partial f_i}{\partial \phi_j}. \end{aligned}$$

Es resultiert zur Berechnung von $c_\phi : \mathbb{R} \rightarrow \mathbb{R}^{s \times \nu}$ das Anfangswertproblem

$$\dot{c}_\phi(t) = (f_c c_\phi + f_\phi)(t, c; \phi), \quad t \in [\bar{t}_1, \bar{t}_k], \quad (7.7)$$

$$c_\phi(\bar{t}_1) = 0. \quad (7.8)$$

Die Komponenten der Jacobimatrix J_g sind

$$(J_g)_{:(j-1)s,:} = \frac{\partial g_{:(j-1)s}}{\partial \phi} = c_\phi(\bar{t}_{j+1}), \quad j = 1, \dots, k-1,$$

mit $c_\phi(\bar{t}_{j+1}) \in \mathbb{R}^{s \times \nu}$ und letztendlich ergibt sich mittels der Lösung des Anfangswertproblems aus (7.7) und (7.8) und den Auswertungen an den Stützstellen \bar{t}_j eine Näherung für ∇G .

Bemerkung 7.3. Zur Auswertung der rechten Seite von (7.7) als zeitliche Ableitung von c_ϕ wird $c(t)$ selbst auch an Zwischenstufen $t \in [\bar{t}_i, \bar{t}_{i+1})$ benötigt. Bei getrenntem Lösen der Anfangswertprobleme aus (7.2), (7.3) und (7.7), (7.8) lassen sich diese Werte nur durch zusätzliche Interpolation bestimmen. Es ist daher sinnvoll beide Anfangswertprobleme zusammen zu lösen und die nötigen Auswertungen von $c(t)$ an den Stufen automatisch mit zu berechnen.

7.2.2 Eine autonome Folge von Anfangswertprobleme

Ein anderer Ansatz ist es, die Parametrisierung des Anfangswertproblems (7.2) in eine Folge von $k - 1$ Anfangswertproblemen der Form

$$\begin{aligned}\dot{c}^{(i)}(t) &= f(t, c^{(i)}; \phi), & t \in [\bar{t}_i, \bar{t}_{i+1}], \\ c^{(i)}(\bar{t}_i) &= C_{i,:}^{(\text{abs})}\end{aligned}\tag{7.9}$$

für $i = 1, \dots, k - 1$ zu überführen. Die Vergleichswerte sind ebenfalls $C_{i+1,:}^{(\text{dgl})} = c^{(i)}(\bar{t}_{i+1})$. Für die Jacobimatrix J_g ist analog zu (7.6) die Lösung der $k - 1$ Anfangswertprobleme

$$\begin{aligned}\dot{c}_\phi^{(i)}(t) &= \left(f_c c_\phi^{(i)} + f_\phi \right) (t, c^{(i)}; \phi), & t \in [\bar{t}_i, \bar{t}_{i+1}], \\ c_\phi^{(i)}(\bar{t}_i) &= 0,\end{aligned}\tag{7.10}$$

für $i = 1, \dots, k - 1$ mit den jeweiligen Approximationen am Intervallende $c_\phi^{(i)}(\bar{t}_{i+1}) \in \mathbb{R}^{s \times \nu}$ erforderlich.

Ein optimaler Parametersatz $\phi^* \in \mathbb{R}^\nu$ mit $G(\phi^*) = 0$ erfüllt die Stetigkeitsbedingungen

$$c^{(i)}(\bar{t}_{i+1}) = C_{i+1,:}^{(\text{abs})} = c^{(i+1)}(\bar{t}_{i+1}), \quad i = 1, \dots, k - 1,$$

und die aus den einzelnen Lösungen $c^{(i)}(t)$ zusammengesetzte Lösung $c(t)$ ist zudem stetig differenzierbar

$$\dot{c}^{(i)}(\bar{t}_{i+1}) = f(\bar{t}_{i+1}, c^{(i)}(\bar{t}_{i+1}); \phi^*) = f(\bar{t}_{i+1}, c^{(i+1)}(\bar{t}_{i+1}); \phi^*) = \dot{c}^{(i+1)}(\bar{t}_{i+1}).$$

Bemerkung 7.4. Ein Unterschied zwischen den Anfangswertproblemen zu den Anfangswerten aus (7.5) und den Anfangswerten aus (7.9) ist die reduzierte Fehlerfortpflanzung bei der Anwendung von (7.9), da für jeden Zeitabschnitt mit dem als korrekt angesehenen Anfangswert gestartet wird. Demgegenüber pflanzt sich mit dem Startwert (7.5) der Fehler durch ungenügende Parameter ϕ von Gitterpunkt zu Gitterpunkt verstärkt fort. Die Fehler zum Zielwert sowie die Ableitungen der Lösung $c(t)$ nach ϕ werden zum Gitterende hin im Normalfall im Vergleich zu vorderen Gitterpunkten verhältnismäßig groß. In Abbildung 7.1 sind die Anwendungen der unterschiedlichen Ansätze anhand eines Beispiels dargestellt.

Bemerkung 7.5. Weiterhin ist es durchaus sinnvoll, die Auswertung der Funktionen g_i aus (7.4) entsprechend dem Kehrwert des Maximaleintrags jeder einzelnen Spalte in \hat{C} zu gewichten

$$g_{i+(j-1)s}(\phi) = (\max_l C_{l,i})^{-1} (c_i(\bar{t}_{j+1}) - C_{j+1,i}), \quad i = 1, \dots, s, \quad j = 1, \dots, k - 1.$$

Bemerkung 7.6. Als Kriterium für ein lokales Minimum ist bei der Parameteroptimierung nur die Konverxität der Funktion in der mutmaßlichen Minimalstelle vorhanden. Also, sofern zusätzlich zu einem verschwindenden Gradienten die Hessematrix der zweiten partiellen Ableitungen der Anpassungsfunktion nach den Parametern ϕ positiv definit ist.

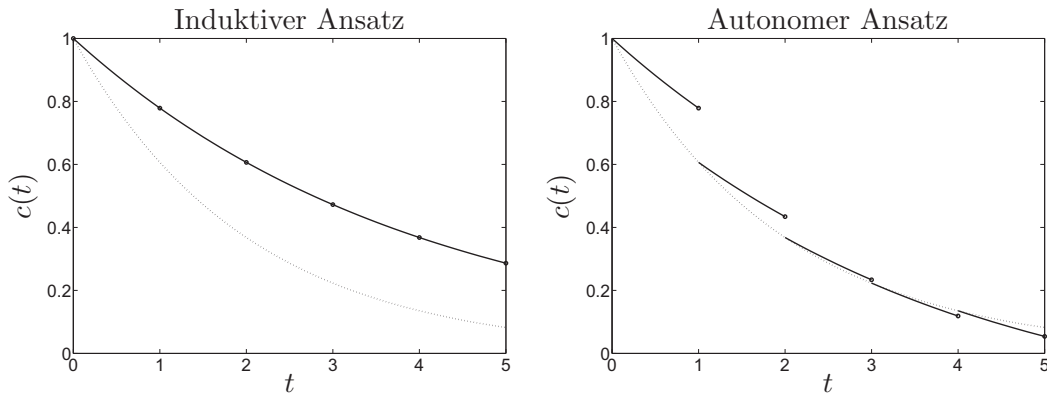


Abbildung 7.1: In den Grafiken sind die Vorgehensweisen der erläuterten Ansätze für eine Beispielrechnung dargestellt. Ausgangspunkt ist die Funktion $c_\phi(t) = \exp\{\phi t\}$ mit $\phi = \phi_0 = -0.5$, der Diskretisierung $\bar{t}_{i+1} = i$, $i = 0, \dots, 5$, und der Differentialgleichung $\dot{c} = \phi c$. Zu dem Parameter $\phi = -0.25$ sind die Lösungen der induktiv aufgebauten Folge von Anfangswertproblemen (links, Anfangswerte aus (7.5)) und die Lösungen der autonom aufgebauten Folge (rechts, Anfangswerte aus (7.9)) dargestellt. Die Fehlerwerte der Funktion $G(\phi)$ ergeben sich aus den Differenzen zwischen den Referenzwerten $c_{\phi_0}(\bar{t}_i)$ und den Auswertungen der Lösungen der Anfangswertprobleme an den Stellen $\bar{t}_2, \dots, \bar{t}_6$.

7.3 Die Implementierung

Die Implementierung der Systemanpassung ist in C realisiert. Zur numerischen Lösung der Anfangswertprobleme werden die FORTRAN Bibliotheken RADAU und DOP853 für steife beziehungsweise nichtsteife Systeme von Hairer und Wanner, [31,32], angewendet. Für die aufgestellte Zielfunktion G lässt sich mit dem Fehlervektor g und dessen Jacobimatrix J_g erneut eine nichtlineare Ausgleichsrechnung wie in Abschnitt 6.3 durchführen. Als Unterschied ist für g die Jacobimatrix zwar nicht analytisch, aber über die Lösung eines oder mehrerer Anfangswertprobleme bestimmbar, siehe (7.7) und (7.10). Alternativ zum Gauß-Newton Verfahren wird in der Praxis erneut die ACM FORTRAN Bibliothek NL2SOL als starkes Werkzeug zur nichtlinearen Ausgleichsrechnung benutzt.

8 Ausblick

In dieser Arbeit wurden multivariate Modelle der Form $A = XY + E$ mit nichtnegativen Faktoren X, Y behandelt. Zu einer gegebenen Matrix A ist eine Zerlegung in $X \in \mathbb{R}_+^{k \times s}$ und $Y \in \mathbb{R}_+^{s \times n}$ mit $\text{rank}(X) = \text{rank}(Y) = s$ unter zusätzlichen Regularisierungen gesucht, so dass das Residuum E in einem geeigneten Sinne klein ist.

Für den störungsfreien Fall $E = 0$ und ohne zusätzliche Regularisierungen stellt sich die Frage nach der Eindeutigkeit der Faktoren. In Bezug darauf sind in Abschnitt 4.4 notwendige und hinreichende Kriterien formuliert. Allerdings sind bereits die notwendigen Voraussetzungen sehr einschränkend und im Normalfall nicht gegeben. In speziellen Fällen sind jedoch Teilfaktoren (einzelne Spalten von X oder Zeilen von Y) in der Faktorisierung eindeutig. Dafür sind die Nichtnegativitätsrestriktionen an die Faktoren und die lineare Struktur möglicher Transformationen entscheidend. Die Eindeutigkeit einiger Teilfaktoren schränkt indirekt auch andere Teilfaktoren ein.

Das gestellte Problem (eine nichtnegative Niedrigrangapproximation zu berechnen) ist somit sowohl überbestimmt (in dem Sinne, dass die bestimmte Lösung nur $A = XY + E$, nicht aber $A = XY$ erfüllt) als auch gleichzeitig unterbestimmt, da die Lösung bis auf triviale Fälle keineswegs eindeutig ist. Ferner kommt noch die unbekannte Anzahl der Teilfaktoren s als diskrete Variable hinzu. Das Problem erweist sich damit als hartnäckig, wegen seines hohen Anwendungspotentials aber auch als sehr lohnend.

Zur Berechnung einer nichtnegativen Niedrigrangapproximation wird eine spezielle Methode angewendet. Ausgangspunkt ist eine Singulärwertzerlegung der Matrix A . Eine Berücksichtigung nur der ersten s Singulärwerte und -vektoren führt auf eine Niedrigrangapproximation. Mittels zweier multiplikativ in die abgeschnittene Singulärwertzerlegung eingefügter Matrizen werden die verbleibenden links- und rechtsseitigen Singulärvektoren zu geeigneten Faktoren X und Y transformiert. Für störungsfreie Daten sind diese Matrizen invers zueinander; das Auftreten von kleinen Störungen ermöglicht den Einsatz des Produkts aus einer Matrix T und ihrer Pseudoinversen

$$A \approx \underbrace{U \Sigma_{:,1:s} T^+}_X \underbrace{T V_{:,1:s}^T}_Y.$$

In beiden Fällen reduziert sich die Konstruktion der Faktoren auf die Berechnung einer Transformation T . Dazu wird ein Optimierungsproblem aufgestellt und gelöst.

Ausgangspunkt der in dieser Arbeit untersuchten Aufgabe ist ein kontinuierlicher Produktansatz $u(x, y) = \sum_{i=1}^s w_i(x) v_i(y)$ mit stetigen Funktionen w_i, v_i . Die Einträge in der zu faktorisierenden Matrix A sind diskrete Auswertungen von u . Ausgehend von diesem Ursprung ist es unter Umständen möglich, Multi-Level-Methoden in Form einer hierarchischen Ausdünnung der Daten einzusetzen. Zu den Grobgitterdaten werden Faktorisierungen berechnet und schrittweise zu Faktorisierungen auf feineren Gittern prolongiert. Dieser Ansatz wird verfolgt, um die Rechenzeit zu reduzieren.

Der kritische Punkt bei der Lösung der schlecht gestellten Faktorisierungsaufgabe ist die Wahl der Regularisierungen und die Spezifizierung ihrer Parameter. Letztendlich kann dieser Punkt nur modellbezogen geklärt werden.

8.1 Die Anwendung der regularisierten Niedrigrangapproximation in der Spektroskopie

„FACTOR ANALYSIS, A MATHEMATICAL TECHNIQUE FOR STUDYING MATRICES OF DATA, IS ONE OF THE MOST POWERFUL METHODS IN THE ARSENAL OF CHEMOMETRICS.“, E. R. Malinowski, [63].

Bei der Analyse eines chemischen Systems mittels spektroskopischer Techniken unterliegt die Absorption der Probe dem idealisierten Gesetz von Lambert-Beer. Für ein Mehrkomponentensystem ist die Gesamtaborption eine Superposition der Einzelabsorptionen. Die Messungen von Absorptionen verschiedener Reinkomponenten und zu diversen Wellenlängen werden in einer Absorptivitätsmatrix \hat{A} (Reinspektren) zusammengefasst. Das Absorptionsspektrum eines Gemischs dieser Stoffe führt auf ein Vektor-Matrix-Produkt; die Absorptionsspektren verschiedener Gemische dieser Komponenten, zusammengefasst in Matrixform, sind das Produkt der Matrizen der Einzelkonzentrationen \hat{C} und der Absorptivitäten \hat{A} . Die Faktoren und damit auch das Produkt sind nichtnegativ. Unter dem Einfluss von Störungen liegt ein Modell $A = \hat{C}\hat{A} + E$ vor.

In dem multivariaten Ansatz ist nur die Matrix A der gemessenen Absorptionen der gesamten Probe zugänglich. Eine Möglichkeit Informationen über das System zu extrahieren, ist dessen Aufschlüsselung in unabhängige Einzelkomponenten. Eine nichtnegative Matrixfaktorisierung von A ist gesucht. Trotz zusätzlicher Systemrestriktionen liegen im Normalfall nicht alle Faktorisierungen in derselben Äquivalenzklasse. Um dennoch tiefgreifende Aufschlüsse über chemische Zusammenhänge zu erlangen, bedarf es einer Selektion unter den möglichen Faktorisierungen mittels einer zusätzlichen Regularisierung der Faktoren.

Computerbasierte Spektroskopietechniken erlauben sehr feine Diskretisierungsgitter und erzeugen große Datenmengen. Bei der Berechnung einer Zerlegung von A ermöglicht dies oft eine stufenweise Ausdünnung der Daten und den Einsatz der in dieser Arbeit entwickelten Multi-Level-Methoden, um die Rechenzeit zu reduzieren.

Für eine vorliegende Datenmatrix sind unter Umständen neben der Nichtnegativität der Faktoren weitere Restriktionen an eine Faktorisierung bekannt. Solche ergeben sich aus Informationen zu dem System, welches der Matrix zugrunde liegt. Unabhängig von einer Regularisierung der Faktoren bieten derartige Restriktionen eine Einschränkung der Interpretationsmöglichkeiten. Von der Menge aller möglichen Zerlegungen wird eine Menge der, für das vorliegende Problem, relevanten abgespalten. Über die Bestimmung einer regularisierten Faktorisierung von A hinaus werden so auch alternative Reinkomponentenzerlegungen und andere Möglichkeiten der Systeminterpretationen aufgezeigt.

8.2 Weiterführende Überlegungen

Zum Abschluss dieser Arbeit sollen einige Perspektiven für die zukünftige Arbeit formuliert werden. Einerseits wird eine auf den Singulärwerten von A basierende Bestimmung des Ranges der Niedrigrangapproximation diskutiert. Andererseits wird eine Fusion des Ansatzes von Kim und Park aus [47] mit dem Ansatz von PCD vorgeschlagen. Zuletzt wird eine Zerlegung mit Nichtnegativitätsrestriktionen in nur einem Faktor betrachtet.

Wahl des Ranges der Niedrigrangapproximation

Der Rang s der Niedrigrangapproximation ist die Anzahl der extrahierten Teilfaktoren; für Analysen von Spektroskopiedaten entspricht s demzufolge der Anzahl der unabhängigen Komponenten. Entscheidende Voraussetzung für Untersuchungen zur Wahl von s sind Kenntnisse

über die Störung E in Form einer Approximation für $\|E\|_2$ oder $\|E\|_F$, der typischen Größe von Einträgen in E oder für zufällige Störungswerte die Kenntnis derer statistischer Verteilung. Derartige Informationen sind anwendungsspezifisch. Für Analysen zur Problematik der Rangwahl der Approximation ist speziell auf Stewart mit [78, 79] verwiesen.

Seien die Singulärwertzerlegungen $U\Sigma V^T = A = XY$ und $\tilde{U}\tilde{\Sigma}\tilde{V}^T = \tilde{A} = XY + E$ der störungsfreien und der gestörten Daten gegeben. Es gelte $\text{rank}(X) = \text{rank}(Y) = s$. Weyl und Mirsky lieferten folgende Eingrenzungen für die Singulärwerte (siehe beispielsweise [78])

$$|\tilde{\sigma}_i - \sigma_i| \leq \|E\|_2, \quad i = 1, \dots, \min\{k, n\}, \quad (8.1)$$

$$\sum_{i=1}^{\min\{k, n\}} (\tilde{\sigma}_i - \sigma_i)^2 \leq \|E\|_F^2. \quad (8.2)$$

Aus (8.1) folgt die Beschränkung

$$s \geq \text{rank}_{\|E\|_2}(\tilde{A})$$

und aus (8.2) die Relation

$$\sum_{i=s+1}^{\min\{k, n\}} \tilde{\sigma}_i^2 \leq \|E\|_F^2. \quad (8.3)$$

Mit Kenntnis von $\|E\|_2$ ist somit $s = \text{rank}_{\|E\|_2}(\tilde{A})$ eine Wahlmöglichkeit. Ist $\|E\|_F$ bekannt, lässt sich s als minimale Wahl, für die (8.3) gilt, anwenden, [79]. Ebenfalls ist eine Entscheidung aufbauend auf den Verhältnissen

$$\tilde{\sigma}_i \tilde{\sigma}_{i+1}^{-1}$$

und einem „Bruch“ in der Folge $\tilde{\sigma}_i$ möglich, was gleichermaßen stark wie die zuvor genannten Entscheidungskriterien von den Einträgen in X, Y und E abhängt.

Bei weiteren Untersuchungen steht in [78] insbesondere der spezielle Fall zufälliger Störungen im Fokus. Für Korrekturen eines idealisierten Modellansatzes sind diese jedoch nur bedingt einsetzbar.

Alternierende Lösung von Ausgleichsproblemen und eine Transformation der Singulärvektoren

In Kapitel 5 sind klassische Zugänge zur Matrixfaktorisierung vorgestellt. In dem von Kim und Park entwickelten Algorithmus aus [47] wird eine Faktorisierung mittels der alternierenden Lösung regularisierter nichtnegativer linearer Ausgleichsprobleme berechnet.

Ein alternativer Lösungsansatz ist eine Vereinigung dieser Idee mit dem in dieser Arbeit genutzten Prinzip der Transformation der ersten z Singulärvektoren von A . Die Optimierung wird also auf niedrigdimensionale Unterräume beschränkt. Werden die Nichtnegativitätsrestriktionen in Straffunktionen eingebettet, so sind alternierend die Minimierungsprobleme

$$\min_{S \in \mathbb{R}^{z \times s}} f_T(S), \quad f_T(S) = \|\Sigma - \Sigma S T\|_F^2 + \sum_{i=1}^{p_S} \gamma_i^2 f_i(U\Sigma_{:,1:z} S), \quad (8.4)$$

$$\min_{T \in \mathbb{R}^{s \times z}} f_S(T), \quad f_S(T) = \|\Sigma - \Sigma S T\|_F^2 + \sum_{i=1}^{p_T} \gamma_i^2 f_i(TV_{:,1:z}^T), \quad (8.5)$$

mit $f_T : \mathbb{R}^{z \times s} \rightarrow \mathbb{R}$ und $f_S : \mathbb{R}^{s \times z} \rightarrow \mathbb{R}$ sowie $U\Sigma V^T$ einer Singulärwertzerlegung von A zu lösen. Anders als in dem Algorithmus von Kim und Park liegen nichtlineare Ausgleichsprobleme vor.

Einerseits verringern sich so die Dimensionen der Variablen gegenüber dem klassischen Ansatz bei $s \leq z \ll \min\{k, n\}$ enorm und andererseits wird die Flexibilität der in PCD eingesetzten Regularisierungen f_i übernommen. Sofern die Regularisierungsfunktionen zusätzlich alle konvex sind, überträgt sich diese Eigenschaft auch auf die Zielfunktionen aus (8.4) und (8.5), da $\|\Sigma - \Sigma ST\|_F^2$ (beim Fixieren von S beziehungsweise T) konvex ist und eine vorgeschaltete lineare Abbildung keinen Einfluss auf die Konvexität der Funktionen f_i hat. Die Problematik möglicher lokaler, aber nicht globaler, Minimalstellen träte in diesem Fall nicht auf. Die Nutzung einer stufenweisen Ausdünnung der Daten für eine Konvergenzbeschleunigung in Form der angewendeten Multi-Level-Methoden ist problemlos einsetzbar.

Erste praktische Anwendungen zu diesem Ansatz sind sehr vielversprechend.

Nichtnegativität in nur einem Faktor

Bezüglich des Ansatzes $A = XY + E$ liegt eine alternative Aufgabenstellung vor, falls nur für einen der Faktoren die Nichtnegativitätsrestriktionen gelten, also beispielsweise $X \in \mathbb{R}_+^{k \times s}$ und $Y \in \mathbb{R}^{s \times n}$. Mit einer solchen Problemstellung der Faktoranalyse beschäftigen sich Kim und Park in [70] und nutzen eine derartige Zerlegung zur Strukturaufklärung am Beispiel von Gesichtserkennungen und Textklassifikationen. Dabei ist in diesen Fällen im Allgemeinen keine Faktorisierung unter kleinen Störungen im Sinne von

$$\sigma_s(XY) \gg \sigma_1(E)$$

gegeben.

Eine klassische Anwendung zur Aufgabenstellung

$$A = XY + E, \quad X \in \mathbb{R}_+^{k \times s}, \quad Y \in \mathbb{R}^{s \times n},$$

mit kleiner Störung E liegt bei der Analyse von Daten aus der Circular dichroismusspektroskopie vor. Die Gesamtabsorption, die Einzelkonzentrationen und die Reinabsorptionen stehen nach dem Gesetz von Lambert-Beer und durch die Superposition der Einzelabsorptionen in einem solchen Zusammenhang. Die negativen Einträge ergeben sich für den Faktor Y aus der Subtraktion der Absorptionsspektren von links- und rechtspolarisiertem Licht. Für den Faktor X der Konzentrationsprofile gelten unverändert Nichtnegativitätsrestriktionen.

Problem dieser Aufgabenstellung ist die gesicherte Nichteindeutigkeit. Jede Transformation mit einer nichtnegativen Matrix T vollen Ranges führt unmittelbar auf eine weitere gültige Niedrigrangapproximation $\tilde{X} = XT$, $\tilde{Y} = T^{-1}Y$. Der Kenntnis externer Systemrestriktionen und der sinnvollen Steuerung der Regularisierungen kommen noch entscheidendere Bedeutungen zu.

Zur Berechnung einer derartigen Faktorisierung stellen Kim und Park in [70] einen Algorithmus der alternierenden Lösung linearer Ausgleichsprobleme vor. Die Bestimmung des Faktors X erfolgt unter Nichtnegativitätsrestriktionen, die von Y nicht. Die Methode arbeitet also analog zu deren in Abschnitt 5.2 vorgestellten Ansatz, eine nichtnegative Matrixfaktorisierung über die alternierende Lösung linearer Ausgleichsprobleme unter Nichtnegativitätsrestriktionen zu bestimmen.

Anhang

A Optische Spektroskopie

In diesem Teil des Anhangs soll kurz die Gewinnung von Messdaten durch die beiden optischen Spektroskopietechniken FT-IR und UV/Vis erläutert werden. Für die Analyse der in den Kapiteln 5 und 6 vorgestellten Methoden zur Berechnung von regularisierten Niedrigrangapproximationen werden in Anhang B vier Matrizen basierend auf Datensätzen zu chemischen Reaktionen faktorisiert. Von den Datensätzen sind drei aus der FT-IR und einer aus der UV/Vis Spektroskopie.

Die Absorptionsspektroskopie beruht allgemein auf der Wechselwirkung elektromagnetischer Strahlung mit Atomen, Molekülen, Ionen und anderen energieaufnehmenden Systemen. Charakterisiert wird die Strahlung üblicherweise über die Frequenz ν , die Wellenlänge λ oder die Wellenzahl $\tilde{\nu}$. Eine relevante Einteilung einzelner Spektralbereiche liefert Tabelle A.1.

Strahlungsart	Frequenz	Wellenlänge	Wellenzahl
Infrarot			
(fern)	0.86 THz – 10.4 THz	350 μm – 29 μm	29 cm^{-1} – 345 cm^{-1}
(mittleres)	10.4 THz – 100 THz	29 μm – 3 μm	345 cm^{-1} – 3 330 cm^{-1}
(nahes)	100 THz – 385 THz	3 μm – 780 nm	3 330 cm^{-1} – 12 800 cm^{-1}
Sichtbares Licht	385 THz – 789 THz	780 nm – 380 nm	12 800 cm^{-1} – 26 300 cm^{-1}
Ultraviolett			
(fern)	789 THz – 1.50 PHz	380 nm – 200 nm	26 300 cm^{-1} – 50 000 cm^{-1}
(nahes)	1.50 PHz – 300 PHz	200 nm – 10 nm	50 000 cm^{-1} – 10^5 cm^{-1}

Tabelle A.1: Ausgewählte Spektralbereiche elektromagnetischer Strahlung, [41].

A.1 Die FT-IR Spektroskopie

Ein Molekül besteht aus zwei oder mehr Atomen, die in einer fest definierten räumlichen Lage zueinander angeordnet sind. Die Abstände der Atomkerne werden dabei durch die Summe aller Kräfte im Gleichgewicht festgelegt. Unter dem Einfluss von IR Strahlung kommt es zu Rotations- und Schwingungsübergängen des Moleküls. Ändert sich das Dipolmoment des Moleküls bei der Schwingung, ist diese IR-aktiv. Die dabei zugrunde liegenden Frequenzen sind charakteristisch für die Molekülbindungen (beziehungsweise für funktionelle Gruppen des Moleküls) und erlauben somit Rückschlüsse auf das Molekül selbst. Für tiefgründigere Erläuterungen ist an dieser Stelle auf [28, 45] verwiesen.

A.1.1 Die Messmethodik

Das Herzstück eines FT-IR Spektrometers ist das Michelson Interferometer. Ursprünglich von dem Physiker Albert Abraham Michelson zur Bestimmung der Lichtgeschwindigkeit entwickelt,

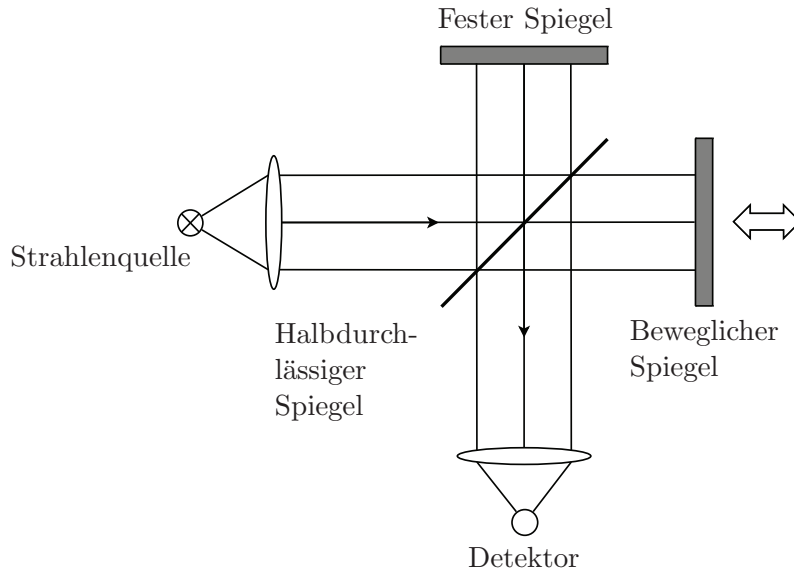


Abbildung A.1: Aufbau eines Michelson Interferometers.

wird die Apparatur heute beispielsweise in der Spektroskopie eingesetzt, um die Strahlungsabsorption einer Probe für einen Wellenlängenbereich zu untersuchen. Durch die Bestrahlung einer Probe mit polychromatischem Licht wird ein Interferogramm erzeugt, aus welchem mittels Fourier-Transformation ein Transmissions- und ein Absorptionsspektrum errechnet wird.

Das Michelson Interferometer (siehe Abbildung A.1) besteht aus einer Strahlenquelle, einem halbdurchlässigen, einem festen und einem beweglichen Spiegel sowie einem Detektor. Nach der Strahlenteilung durch den halbdurchlässigen Spiegel wird der eine Strahl vom festen Spiegel und der andere von dem beweglichen Spiegel zum Strahlenteiler reflektiert. Beide Strahlen werden unter Verlusten wieder vereinigt, zur Interferenz gebracht und der Detektor nimmt die Intensität des überlagerten Strahls auf. Entscheidend ist dabei der Wegunterschied, den das Licht über den beweglichen Spiegel im Vergleich zu dem Weg über den festen Spiegel besitzt. Für monochromatisches Licht der Wellenlänge λ , der Lichtintensität I_0 vor der Strahlenteilung und einen Wegunterschied x der geteilten Lichtstrahlen ergibt sich die Intensität

$$I(x) = \frac{1}{2}I_0(1 + \cos(\frac{2\pi}{\lambda}x)).$$

Der Gangunterschied 0 bedeutet demnach konstruktive Interferenz, $x = \lambda/2$ führt auf destruktive Interferenz. Der entscheidende Vorteil dieser Spektroskopiemethode ist die Verwendung polychromatischen Lichts. Die am Detektor gemessene Lichtintensität $I(x)$ und die Energiedichte $B(\tilde{\nu})$ bei einer bestimmten Wellenzahl $\tilde{\nu} = \lambda^{-1}$ ergeben zusammen das Cosinus-Fourier-Transformations Paar

$$\begin{aligned} I(x) &= \int_{-\infty}^{\infty} B(\tilde{\nu}) \cos(2\pi\tilde{\nu}x) d\tilde{\nu}, \\ B(\tilde{\nu}) &= \int_{-\infty}^{\infty} I(x) \cos(2\pi\tilde{\nu}x) dx. \end{aligned} \tag{A.1}$$

Mittels des Interferogramms $I(x)$ lässt sich über (A.1) die Energiedichte $B(\tilde{\nu})$ und weiter die Absorption $A(\tilde{\nu})$ beziehungsweise $A(\lambda)$ bestimmen. Nähere Erläuterungen bezüglich des Übergangs von der Orts-Intensitäts-Abhängigkeit zur Wellenzahl-Energie-Abhängigkeit finden sich in [24, 46].

Wird nun vor dem Detektor eine Küvette mit einer Probe platziert, so lässt sich ein Interferogramm zu der Probe aufnehmen und dieses mittels Fourier-Transformation in ein Transmissions- sowie weiter in ein Absorptionsspektrum umwandeln.

Einen quantitativen Zusammenhang zwischen Absorption und Konzentration einer Probe beschreibt das Gesetz von Lambert-Beer mit der Beziehung aus (2.10). Zusammen mit der Superposition für mehrere Substanzen folgt so die Überleitung zum mathematischen Teil dieser Arbeit.

A.1.2 Nichtlinearitätskorrekturen und Fehlerterme

Idealisiert wird bei der Analyse spektroskopischer Daten ein Produkt $A = \hat{C}\hat{A}$ betrachtet. Unter dem Einfluss von Störungen wird ein Modell der Form

$$A = \hat{C}\hat{A} + E, \quad E \in \mathbb{R}^{k \times n},$$

untersucht. Die Störungen setzen sich in praktischen Anwendungen unter anderem aus folgenden Ursachen zusammen:

1. Der lineare Zusammenhang des Gesetzes von Lambert-Beer in beiden Faktoren \hat{C}, \hat{A} ist nur eine idealisierte Näherung.
2. Messtechnische Probleme, unter anderem durch Ablagerungen am Fenstermaterial der Transmissionszelle, nicht ideale Durchmischung, nicht ideale Durchpumpvorgänge sowie Temperaturschwankungen.
3. Bei der Vorbehandlung der Rohdaten erfolgt die Subtraktion eines Hintergrundspektrums. Der Ablauf eines Versuchs in einem Lösungsmittel liefert Beiträge aller in der Küvette vorhandenen Komponenten, also auch vom Lösungsmittel. Während des Ablaufs einer Reaktion bleibt dessen Konzentrationen \hat{c}_{LM} zeitlich konstant und bildet eine zusätzliche unabhängige Komponente mit der Absorption $\hat{a}_{LM} \in \mathbb{R}^n$. In manchen Anwendungen ist $\hat{c}_{LM}\hat{a}_{LM}$ ein sehr stark dominierender Anteil der einzelnen Spektren; dies erschwert die Faktorisierung der Daten in dem Sinne, dass unter Umständen die Absorptionsanteile der anderen Substanzen Wertemäßig im Bereich der Störungen des Hintergrundspektrums liegen. Um trotzdem eine geeignete Zerlegung zu generieren, wird die Vereinfachung

$$A' = A - (\hat{c}_{LM}, \hat{c}_{LM}, \dots, \hat{c}_{LM})^T \hat{a}_{LM} l$$

angewendet, da sich $\hat{c}_{LM}\hat{a}_{LM}l$ als eine Einzelmessung vor Beginn der Reaktion bestimmen lässt. Die unter Punkt 2 genannten Umstände erzeugen in der Praxis Ungenauigkeiten, so dass teilweise nach Subtraktion des Hintergrunds eine theoretisch nicht mögliche Matrix A' mit

$$\exists I \subset \{1, \dots, k\}, J \subset \{1, \dots, s\}, I, J \neq \emptyset : \quad A'_{i,j} < 0, \quad \forall i \in I, j \in J$$

vorliegt.

4. Häufig ist in Spektrenserien ein Drift der Nulllinie zu beobachten. Dieser ist nicht nur zeit-, sondern auch leicht wellenlängenabhängig und erschwert die Faktorisierung stark. Im Extremfall beinhaltet jedes Spektrum $A_{i,:}$, $i = 1, \dots, k$, eine eigene, veränderte Grundlinie und so werden weitere „künstliche“ Komponenten erzeugt. Um für die spätere Faktorisierung diese Problematik auszuschließen, wird eine Grundlinienkorrektur vorgenommen, siehe beispielsweise Abbildung A.2.

A.2 Die UV/Vis Spektroskopie

Die Strahlung im UV- und im sichtbaren Wellenlängenbereich ist im Vergleich zur IR-Strahlung energiereicher. Die Absorption des Lichts durch die chemische Probe ist dabei hauptsächlich auf

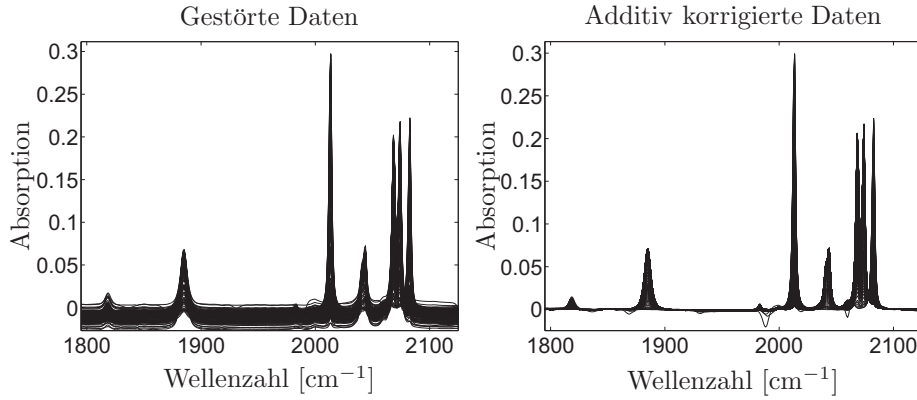


Abbildung A.2: Links: Ausgewählte Absorptionsspektren eines später in Anhang B.4 vorgestellten Versuchs. Eine Subtraktion des Hintergrundspektrums ist bereits erfolgt, was die negativen Einträge erklärt. Zu erkennen ist ein zeitlicher Drift der Nulllinie. Rechts: Additiv korrigierte Daten. Dabei ist jede Linie $A_{i,j}$, $i = 1, \dots, k$, einzeln durch ein Polynom vom Grad 1 korrigiert. Die negativen Einträge aus der Störung bleiben bestehen.

eine Elektronenanregung zurückzuführen, wodurch die Elektronen auf ein höheres Energieniveau gehoben werden. Das Gesetz von Lambert-Beer und die Superposition der Einzelabsorptionen gelten analog, weshalb auch hierfür eine Systemaufklärung mittels nichtnegativer Matrixfaktorisierung der Messmatrix möglich ist.

A.2.1 Die Messmethodik

Für die UV/Vis-Spektroskopie stehen verschiedene Messmethoden zur Verfügung, beispielsweise das klassische Spektrometer oder ein Diodenarray.

Klassisches Spektrometer

Bei diesem ursprünglichen Messgerät werden mittels eines Monochromators die Absorptionen zu den Wellenlängen einzeln vermessen und nicht die des gesamten Wellenlängenbereichs gleichzeitig. Der Vorteil dieser Methodik ist ein sehr gutes Signal-Rausch-Verhältnis, der Nachteil die Dauer eines Durchlaufs aller Wellenlängen, welcher bei der Analyse einer chemischen Reaktion unter Umständen nicht infinitesimal klein im Vergleich zu deren Dynamik ist. Es ergibt sich eine Datenmatrix

$$A_{i,j} \approx u(\bar{t}_i + \Delta t \frac{j-1}{n-1}, \bar{x}_j), \quad i = 1, \dots, k, \quad j = 1, \dots, n,$$

mit $\Delta t > 0$ dem Zeitabstand zwischen dem Abtasten mit \bar{x}_1 und dem Abtasten zu \bar{x}_n . Die Annahme einer (im Vergleich zum Produkt $\hat{C}\hat{A}$) kleinen Störung E in dem multivariaten Ansatz $A = \hat{C}\hat{A} + E$ ist so schwer zu halten.

Diodenarray

Im Gegensatz zum Monochromator wird mit dem Diodenarray das gesamte Spektrum des emittierten Lichts der Strahlungsquelle gleichzeitig gemessen. Die Anzahl der Photodioden, zuständig für die Transmissionsaufnahme, ist dabei vom Gerät abhängig. Die Spektren des in Anhang B.3 betrachteten Versuchs sind durch ein Gerät mit 256 Photodioden aufgenommen.

B Numerische Ergebnisse der regularisierten Niedrigrangapproximation

Die in Anhang A beschriebenen Spektroskopietechniken IR und UV/Vis liefern Daten, die auf dem Gesetz von Lambert-Beer basieren. Zur Analyse der zugrunde liegenden Probe kann eine regularisierte Niedrigrangapproximation der Datenmatrix beitragen. Zu 5 Fallstudien werden in diesem Teil des Anhangs numerische Ergebnisse zur Berechnung von regularisierten Niedrigrangapproximationen präsentiert. Eingesetzt wird dazu die neu entwickelte Methode PCD aus Kapitel 6. Weiter werden ebenfalls die in Kapitel 5 vorgestellten klassischen Algorithmen zur Matrixfaktorisierung auf die Daten angewendet und die Ergebnisse kritisch verglichen.

Insbesondere wird die Funktionsweise der Innovationen des Algorithmus PCD genauer analysiert. Ebenso erfolgt eine Betrachtung der Lösungseinschränkung unter Systemrestriktionen.

B.1 Ein Modellproblem

Für eine einführende Analyse ist es sinnvoll ein Modellproblem zu entwickeln. Als ein solches eignet sich das folgende 2-Komponentensystem, veröffentlicht in [64]: „*Pure component spectral recovery and constrained matrix factorizations: Concepts and applications*“.

Zu einem Verschiebungsparameter $\gamma \in \mathbb{R}$ werden die Faktoren

$$\hat{a}_1(x) = \frac{3}{2} \exp\left(-\frac{(x-250)^2}{100}\right) + 3 \exp\left(-\frac{(x-200)^2}{100}\right) + \frac{3}{2} \exp\left(-\frac{(x-150)^2}{100}\right), \quad (\text{B.1a})$$

$$\hat{a}_2(x) = 2 \exp\left(-\frac{(x-50)^2}{30\,000}\right) + 1.3 \exp\left(-\frac{(x-200-\gamma)^2}{1000}\right), \quad (\text{B.1b})$$

$$\hat{c}_1(t) = 1 - \hat{c}_2(t), \quad (\text{B.1c})$$

$$\hat{c}_2(t) = \frac{\exp(0.1t) - 1}{\exp(0.1t) + 10} \quad (\text{B.1d})$$

betrachtet. Für die Überführung zur diskreten Faktorisierungsaufgabe ist $\gamma = 20$ sowie die Diskretisierung

$$\bar{t} = (0, 1, 2, \dots, 100)^T \in \mathbb{R}^{101}, \quad \bar{x} = (0, 1, 2, \dots, 500)^T \in \mathbb{R}^{501}$$

gewählt. Es ergibt sich die Produktmatrix $A \in \mathbb{R}^{101 \times 501}$,

$$A_{i,j} = \hat{c}_1(\bar{t}_i) \hat{a}_1(\bar{x}_j) + \hat{c}_2(\bar{t}_i) \hat{a}_2(\bar{x}_j), \quad i = 1, \dots, 101, \quad j = 1, \dots, 501.$$

Aufgabe der nichtnegativen Matrixfaktorisierung

Ziel ist es, für die vorliegende Matrix A eine nichtnegative Matrixfaktorisierung zu bestimmen, also anhand von A auf zwei Matrizen $\hat{C} \in \mathbb{R}_+^{101 \times 2}$ und $\hat{A} \in \mathbb{R}_+^{2 \times 501}$ zu schließen, sodass gilt

$$\hat{C}_{i,j} = \hat{c}_j(\bar{t}_i), \quad \hat{A}_{j,l} = \hat{a}_j(\bar{x}_l), \quad i = 1, \dots, 101, \quad j = 1, 2, \quad l = 1, \dots, 501.$$

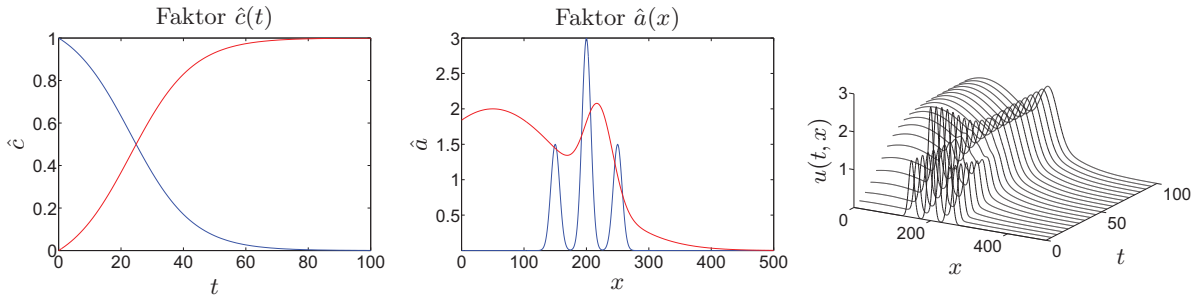


Abbildung B.1: Darstellung der Teilfaktoren $\hat{c}(t)$ und $\hat{a}(x)$ sowie $u(t, x) = \hat{c}_1(t)\hat{a}_1(x) + \hat{c}_2(t)\hat{a}_2(x)$ zu dem beschriebenen 2-Komponentensystem mit $\gamma = 20$ (ausgedünnt auf jede fünfte Linie). Blau: (\hat{c}_1, \hat{a}_1) , rot: (\hat{c}_2, \hat{a}_2) .

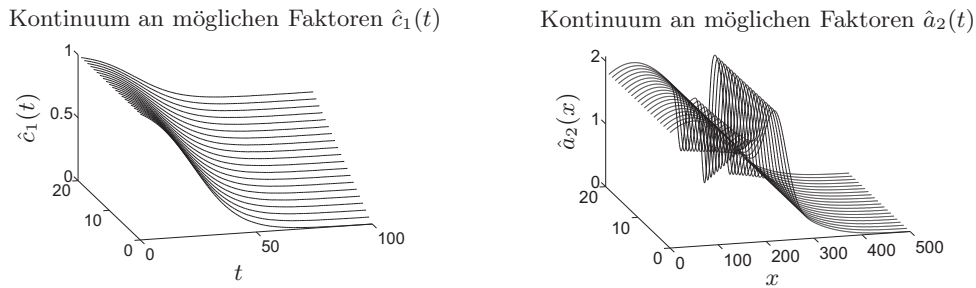


Abbildung B.2: Die möglichen, zueinander komplementären Faktoren $\hat{c}_1(t)$ und $\hat{a}_2(x)$ für eine nichtnegative Matrixfaktorisierung des Datensatzes.

B.1.1 Systemfakten und Mehrdeutigkeit der Faktorisierung

Zu dem System selbst und der Eindeutigkeitsfrage bezüglich der Faktorisierung dessen Datenmatrix ist Folgendes bekannt.

Systemfakten:

- Modellproblem,
- Veröffentlicht in *Journal of Chemometrics*: „Pure component spectral recovery and constrained matrix factorizations: Concepts and applications“ [64],
- Betrachteter Ausschnitt: $t \in [0, 100]$, $x \in [0, 500]$,
- Dimensionen: $k = 101$, $n = 501$, rekonstruierbare Faktorenpaare $s = 2$,
- Abbildung B.1.

Aussagen bezüglich der Mehrdeutigkeit der Faktorisierung:

- Es gibt keine eindeutig bestimmte Äquivalenzklasse an Faktoren.
- Es gilt $\hat{a}(0) \approx (0, 1.84)^T$, $\hat{c}(0) = (1, 0)^T$, $\hat{c}(100) \approx (5E - 5, 1)^T$. Damit ist Satz 3.15 abgesehen von Störungen $\varepsilon \approx 5E - 5$ anwendbar und es ist $\hat{c}_2(t)$ ohne Skalierungsfaktor und im Sinne von Störungen eindeutig in der Faktorisierung.
- Weiter ist Satz 3.19 anwendbar und $\hat{a}_1(x)$ ist bis auf die Skalierung eindeutig.
- Es ergeben sich nach den Sätzen 3.25 und 3.28 für $\hat{c}_1(t)$ und $\hat{a}_2(x)$ jeweils eindimensionale Kontinua an möglichen Faktoren, siehe Abbildung B.2.

B.1.2 Analyse der Daten

Vor der Lösung der Faktorisierungsaufgabe erfolgt eine kurze Analyse der Daten. Dazu werden einige Kenngrößen zur Singulärwertzerlegung des Datensatzes bestimmt.

i	σ_i	$\ A\ _F^{-2} \sum_{j=i+1}^k \sigma_j^2$	$\frac{1}{\ U_{:,i}\ _1} \sum_{j=1}^k U_{j,i}$	$\left\ \frac{d^2 U_{:,i}}{dt^2} \right\ _2$	$\frac{1}{\ V_{:,i}\ _1} \sum_{j=1}^n V_{j,i}$	$\left\ \frac{d^2 V_{:,i}}{dx^2} \right\ _2$
1	1.529 E+02	1.871 E-01	-1.000 E+00	5.938 E-04	-1.000 E+00	8.546 E-03
2	7.336 E+01	0.000 E+00	3.573 E-01	1.687 E-03	-4.429 E-01	1.727 E-02
3	1.092 E-13	0.000 E+00	3.916 E-16	1.744 E+00	3.060 E-02	8.673 E-01

Tabelle B.1: Einige numerisch bestimmte Kenngrößen der in diesem Abschnitt untersuchten Matrix (Singulärwertzerlegung $U\Sigma V^T = A$ und σ_i die Diagonaleinträge von Σ).

Vor dem Hintergrund eines ungestörten Modellproblems erfolgt für die Bestimmung einer nicht-negativen Matrixfaktorisierung durch die verschiedenen Algorithmen die Wahl $s = 2$.

Für den BTEM- und den PCD-Algorithmus wird in (5.15) beziehungsweise (6.2) die Dimension $z = 2$ gesetzt.

B.1.3 Die Ergebnisse der regularisierten Niedrigrangapproximation

Auf die Datenmatrix werden die in den Kapiteln 5 und 6 beschriebenen Verfahren zur Bestimmung einer regularisierten Faktorisierung angewendet. Die Ergebnisse sind in Abbildung B.3 dargestellt.

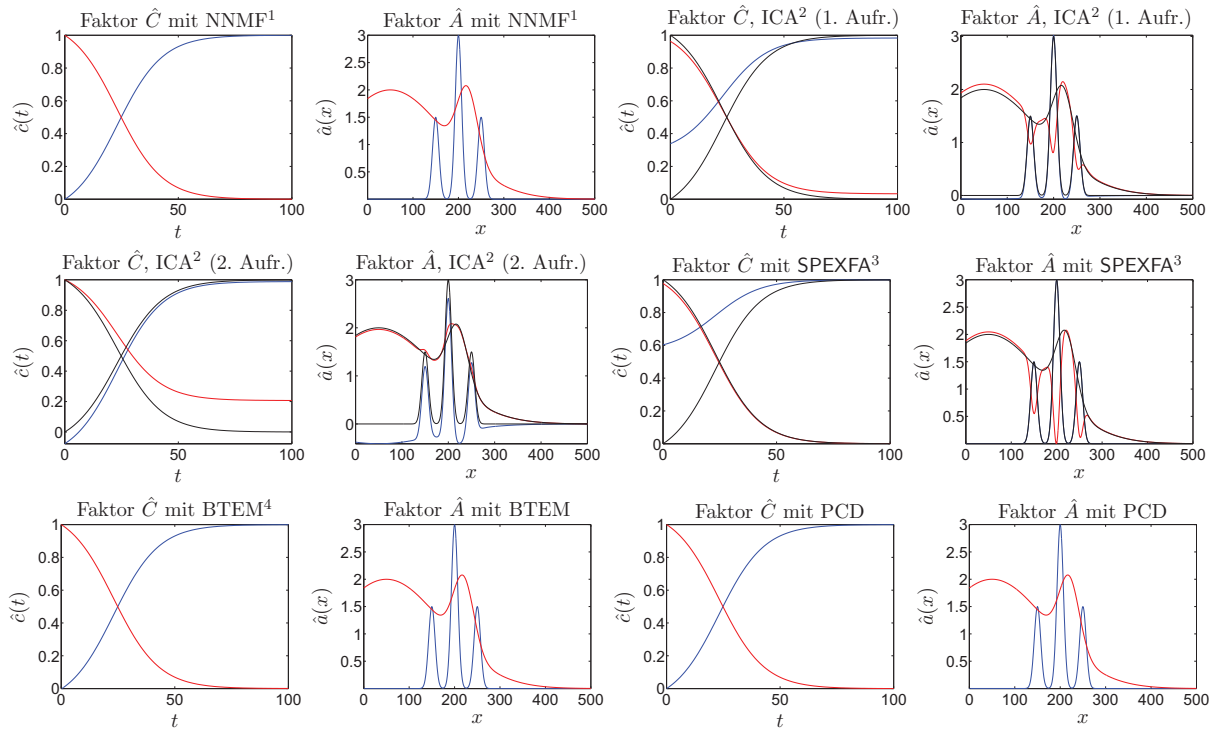


Abbildung B.3: Die Ergebnisse der Algorithmen aus Kapitel 5 sowie PCD für das Modellproblem. Die schwarzen Linien gehören dabei zu den Ausgangskurven aus (B.1).

¹Berechnet mit der MATLAB Routine `nmfans_comb` von Kim und Park.

²Berechnet mit dem MATLAB Code `FastICA_25`, bereitgestellt unter <http://www.cis.hut.fi/projects/ica/fastica/>.

³Berechnet mit der MATLAB Routine `SPEXFA` aus der begleitenden CD von [63].

⁴Im Nachhinein durch lineare Ausgleichsrechnung auf den Zielbereichen bestimmt.

Skalierung in Abbildung B.3:

- Die Ergebnisse der SPEXFA-Methode sind mit der Skalierung der Originaldaten bezüglich der Maximumsnorm angepasst, also $\hat{A}_{1,:}$ und $\hat{A}_{2,:}$ auf $\hat{a}_1(\bar{x})$ und $\hat{a}_2(\bar{x})$.
- Die Ergebnisse der anderen Methoden sind bezüglich der euklidischen Norm angepasst.

Anmerkungen zu den Ergebnissen und Abbildung B.3:

1. Das Ergebnis des Algorithmus ICA ist wenig aussagekräftig, da die Nichtnegativität der Faktoren nicht beachtet wird.
2. Der ICA-Algorithmus gibt bei mehrfacher Anwendung in der Regel unterschiedliche Ergebnisse aus. Es sind die Faktoren von zwei verschiedenen Zerlegungen dargestellt.
3. Eine Anwendung des NNMF-Algorithmus von Kim und Park auf A^T liefert Ergebnisse wie SPEXFA.
4. Der Parametersatz des BTEM-Algorithmus (Zielbereiche (ZB) aus Definition 5.6 und die Parameter aus (5.17)):
 - Linie 1 (blau): ZB: [195, 205]; $\delta_{Sh} = 0$, $\delta_{Der} = \delta_{Int} = 1$.
 - Linie 2 (rot): ZB: [48, 54]; $\delta_{Sh} = \delta_{Int} = 0$, $\delta_{Der} = 1$.
5. Der Parametersatz des PCD-Algorithmus in (6.11): $\gamma_1 = 4$, $\gamma_2 = 4$, $\gamma_3 = 1$, $\gamma_5 = 0.4$ sowie $\gamma_i = 0$, $i \in \{4, 6, 7, 8, 9, 10, 11, 12, 13\}$.
6. Das Problem besitzt unendlich viele Äquivalenzklassen an Zerlegungen, insofern kann es kein „richtig“ und kein „falsch“ geben, solange die Nichtnegativitäts- und Rekonstruktionsforderungen beachtet sind.

B.1.4 Die Analyse der Ergebnisse

Für das Modellproblem lassen sich die Ergebnisse und die Ausgangsdaten mit Hilfe der Größen

$$\varepsilon_i = \|\hat{A}_{i,:}^{(\text{Orig})} - \hat{A}_{i,:}^{(\text{Alg})}\|_2, \quad i = 1, 2, \quad (\text{B.2})$$

$$\epsilon_i = \|\hat{C}_{:,i}^{(\text{Orig})} - \hat{C}_{:,i}^{(\text{Alg})}\|_2, \quad i = 1, 2, \quad (\text{B.3})$$

vergleichen, siehe Tabelle B.2.

	NNMF ¹	ICA ² 1. Aufruf	ICA ² 2. Aufruf	SPEXFA ³	BTEM ⁴	PCD
ε_1	8.61 E−03	8.95 E−01	5.86 E+00	1.94 E−14	7.95 E−06	1.32 E−02
ε_2	1.38 E−02	4.27 E+00	8.95 E−01	7.76 E+00	3.53 E−02	3.82 E−06
ϵ_1	4.62 E−03	1.43 E+00	3.64 E−01	2.59 E+00	(1.25 E−02)	2.51 E−03
ϵ_2	2.64 E−03	2.76 E−01	1.68 E+00	9.66 E−02	(1.03 E−03)	4.15 E−03

Tabelle B.2: Die Tabelle zeigt die Fehlerwerte $\varepsilon_1, \varepsilon_2$ und ϵ_1, ϵ_2 aus (B.2) und (B.3) bezüglich der Skalierung wie in Abbildung B.3.

¹Berechnet mit der MATLAB Routine `nmfanls_comb` von Kim und Park.

²Berechnet mit dem MATLAB Code `FastICA_25`, bereitgestellt unter <http://www.cis.hut.fi/projects/ica/fastica/>.

³Berechnet mit der MATLAB Routine `SPEXFA` aus der begleitenden CD von [63].

⁴Zur Berechnung der Größen ϵ_1, ϵ_2 wurde $\hat{C}_i^{(\text{BTEM})}$ extern über eine lineare Ausgleichsrechnung bestimmt.

B.2 Die Hydroformylierung von 3,3-Dimethyl-1-buten

Im Zuge der Forschungsarbeiten am Leibniz Institut für Katalyse in Rostock und dem Promotionsprojekt von Dipl. Chem. Christoph Kubis entstand folgende Zusammenarbeit zur Hydroformylierung von 3,3-Dimethyl-1-buten: „*A comparative in situ HP-FTIR spectroscopic study of bi- and monodentate phosphite-modified hydroformylation*“ [53]. Die Analyse basiert dabei auf den Ergebnissen des Algorithmus PCD.

In diesem Teilabschnitt soll ein Vergleich der Ergebnisse der regularisierten Niedrigrangapproximation mit denen anderer Methoden gezogen werden.

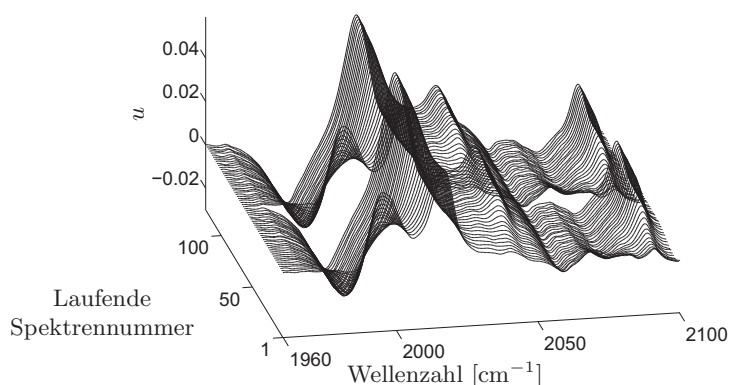


Abbildung B.4: Ausschnitt des Systems der Hydroformylierung von 3,3-Dimethyl-1-buten, ausgedünnt auf jedes zweite Spektrum, zwei Versuchsdurchläufe aneinandergereiht.

B.2.1 Systemfakten und Mehrdeutigkeit der Faktorisierung

Zu dem System selbst und der Eindeutigkeitsfrage bezüglich der Faktorisierung dessen Datenmatrix ist Folgendes bekannt.

Systemfakten:

- Dr. D. Selent, Dipl. Chem. Chr. Kubis: Hydroformylierung von 3,3-Dimethyl-1-buten zu 4,4-Dimethylpentanal und 2,3,3-Trimethylbutanal,
- Veröffentlicht in *ChemCatChem*: „*A Comparative In Situ HP-FTIR Spectroscopic Study of Bi- and Monodentate Phosphite-Modified Hydroformylation*“ [53],
- Reaktionstemperatur: 70° C, Druck: 20 bar Synthesegas ($CO : H_2 = 1 : 1$), Olefinkonzentration: 0.9 mol l^{-1} , Rhodiumkonzentration: $3 \cdot 10^{-4} \text{ mol l}^{-1}$, Lösungsmittel: Cyclohexan,
- Spektroskopieart: FT-IR,
- Zwei Versuchsdurchläufe unter gleichen Bedingungen (64 und 61 Spektrenaufnahmen), für die Niedrigrangapproximation zu einer Matrix (125 Spektren) aneinandergereiht,
- Betrachteter Ausschnitt: $t \in [2.01 \text{ min}, 42.11 \text{ min}]$ beziehungsweise $t \in [2.01 \text{ min}, 40.20 \text{ min}]$ und $\tilde{\nu} \in [1960 \text{ cm}^{-1}, 2100 \text{ cm}^{-1}]$,
- Dimensionen: $k = 125$, $n = 292$, rekonstruierbare Komponenten im betrachteten Spektralbereich: $s = 3$,
- Vorherrschende Komponenten im betrachteten Spektralbereich: Katalysatoren Acylkomplex und Hydridokomplex sowie Restabsorptionen des aus den Daten bereits extrahierten Olefins und eine Störung durch den Hintergrund (hauptsächlich Lösungsmittel),

- Nicht zu vernachlässigende Beiträge durch die gestörte Subtraktion des Olefinspektrums sowie des Hintergrundspektrums: Es gibt i_0 mit $c_{i_0}(t) \not\equiv 0$, diese Störung wird über (6.7) eingebunden,
- Abbildung B.4.

Aussagen bezüglich der Mehrdeutigkeit der Faktorisierung:

- Es liegt keine Eindeutigkeit der Niedrigrangapproximation vor.
- Annahme: $\exists i_1 : c_{i_1}(\bar{t}_{64}) = \hat{C}_{64,i_1} = c_{i_1}(\bar{t}_{125}) = \hat{C}_{125,i_1} = 0$.

B.2.2 Analyse der Daten und Parameterwahl

Eine Analyse des Datensatzes mittels Singulärwertzerlegung ergibt die in der folgenden Tabelle B.3 aufgelisteten Werte. In Abbildung B.5 sind die links- und rechtsseitigen Singulärvektoren dargestellt.

i	σ_i	$\ A\ _F^{-2} \sum_{j=i+1}^k \sigma_j^2$	$\frac{1}{\ U_{:,i}\ _1} \sum_{j=1}^k U_{j,i}$	$\left\ \frac{d^2 U_{:,i}}{dt^2} \right\ _2$	$\frac{1}{\ V_{:,i}\ _1} \sum_{j=1}^n V_{j,i}$	$\left\ \frac{d^2 V_{:,i}}{d\tilde{\nu}^2} \right\ _2$
1	2.780 E+00	9.223 E-02	-1.000 E+00	8.849 E-02	-2.250 E-01	1.161 E-02
2	8.718 E-01	2.980 E-03	-2.558 E-01	3.982 E-01	-3.660 E-01	1.439 E-02
3	1.483 E-01	3.974 E-04	1.335 E-02	3.295 E-01	-9.026 E-01	1.464 E-02
4	4.708 E-02	1.370 E-04	-1.540 E-02	5.970 E-01	1.105 E-01	3.813 E-02
5	1.772 E-02	1.002 E-04	6.113 E-03	1.732 E+00	6.841 E-01	4.808 E-02
6	1.089 E-02	8.625 E-05	5.828 E-03	1.942 E+00	-3.843 E-01	9.367 E-02
7	6.797 E-03	8.082 E-05	3.325 E-03	1.944 E+00	-1.391 E-02	1.395 E-01
8	6.462 E-03	7.592 E-05	-8.503 E-04	2.132 E+00	3.362 E-02	1.668 E-01
9	5.958 E-03	7.175 E-05	-9.882 E-04	2.454 E+00	-5.123 E-02	2.214 E-01
10	5.546 E-03	6.813 E-05	6.670 E-04	2.435 E+00	-1.661 E-03	1.856 E-01
20	4.059 E-03	4.311 E-05	-4.749 E-04	2.431 E+00	3.250 E-02	2.388 E-01

Tabelle B.3: Analyse des in diesem Abschnitt untersuchten Datensatzes mittels Singulärwertzerlegung.

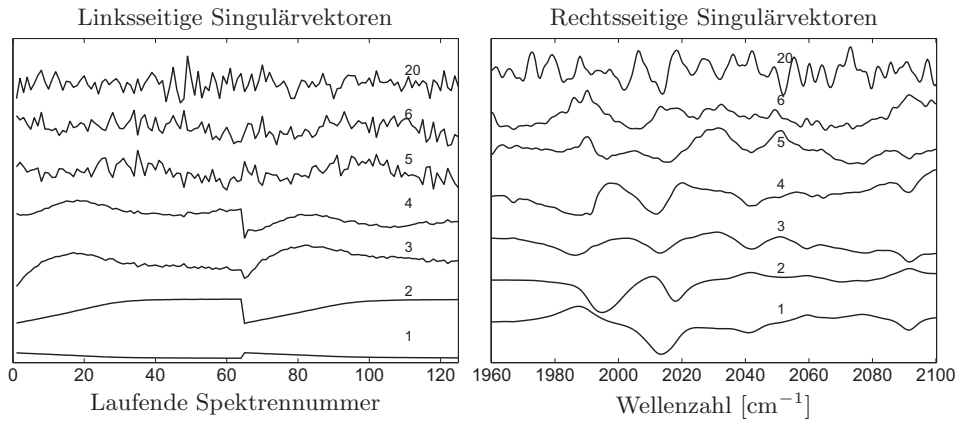


Abbildung B.5: Darstellung der links- und rechtsseitigen Singulärvektoren aus der Singulärwertzerlegung der Datenmatrix.

Für die Anwendung der Algorithmen zur Bestimmung einer Niedrigrangapproximation des zugrunde liegenden Datensatzes erfolgt anhand von Tabelle B.3 und Abbildung B.5, der Glattheit der Singulärvektoren sowie deren Gehalt an Informationen die Wahl $s = 3$.

Der zusätzliche Parameter z wird für den BTEM- und den PCD-Algorithmus als $z = 5$ gewählt. Dabei richtet sich das Augenmerk für diese Wahl hauptsächlich auf die rechtsseitigen Singulärvektoren.

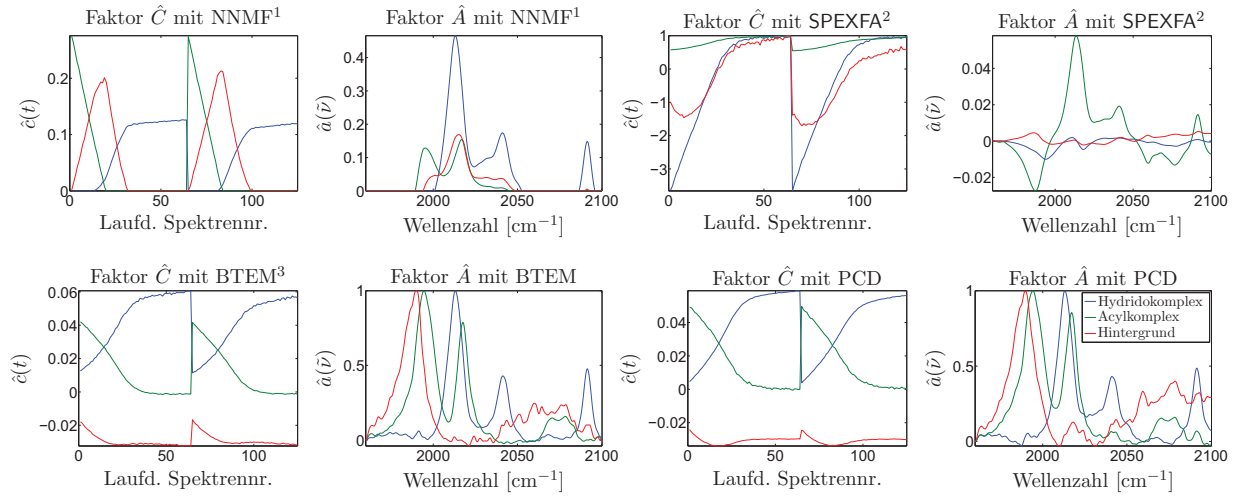


Abbildung B.6: Niedrigrangapproximationen des Datensatzes.

B.2.3 Die Ergebnisse der regulierten Niedrigrangapproximation

Die Anwendung der in dieser Arbeit beschriebenen Methoden zur Bestimmung einer regulierten Niedrigrangapproximation der Datenmatrix ergibt unterschiedliche Ergebnisse, siehe Abbildung B.6.

Aufgrund der Störungen und damit verbunden den teilweise negativen Einträgen der Ausgangsmatrix müssen die Ergebnisse auch unterschiedlich bewertet werden. So korrigiert der eingesetzte klassische NNMF-Algorithmus von Kim und Park alle negativen Einträge vor Beginn der Rechnung auf Null.

Anmerkungen zu den Ergebnissen und Abbildung B.6:

1. Der NNMF-Algorithmus setzt in den Ausgangsdaten alle negativen Einträge auf Null.
2. Der Parametersatz des BTEM-Algorithmus (Zielbereiche (ZB) aus Definition 5.6 und die Parameter aus (5.17)):
 - Linie 1 (blau): ZB: [2038, 2044]; $\delta_{Sh} = 0$, $\delta_{Der} = \delta_{Int} = 1$,
 - Linie 2 (grün): ZB: [1993, 2000]; $\delta_{Sh} = 0$, $\delta_{Der} = \delta_{Int} = 1$,
 - Linie 3 (rot): ZB: [1975, 1990]; $\delta_{Sh} = 0$, $\delta_{Der} = \delta_{Int} = 1$.
3. Die Parameterwahl des PCD-Algorithmus in (6.11): $\gamma_1 = 100$, $\gamma_2 = 100$, $\gamma_3 = 100$, $\gamma_5 = 0.15$, $\gamma_6 = 1 \text{ E-}3$, $\gamma_{11} = 5$ mit $I = \{64, 125\}$ und $c^{(s)} = (1, 1)^T$ aus (6.10) sowie $\gamma_i = 0$, $i \in \{4, 7, 8, 9, 10, 12, 13\}$.
4. Die Bestimmung einer Faktorisierung mit isolierten Banden ist durch die starke Überlagerung der Banden ($\tilde{\nu} = 1990 \text{ cm}^{-1}$ (rot) und 1994.3 cm^{-1} (grün) sowie $\tilde{\nu} = 2013 \text{ cm}^{-1}$ (blau) und 2017.3 cm^{-1} (grün), Ergebnis von PCD) erschwert.

¹Berechnet mit der MATLAB Routine `nmfans_comb` von Kim und Park.

²Berechnet mit der MATLAB Routine `SPEXFA` aus der begleitenden CD von [63].

³Im Nachhinein durch lineare Ausgleichsrechnung auf den Zielbereichen bestimmt.

B.2.4 Die Analyse der Ergebnisse

Eine objektive Bewertung der Ergebnisse ist aufgrund der Mehrdeutigkeit der Niedrigrangapproximation nicht möglich. Jedoch bieten die Größen

$$\varepsilon_i = \sum_{j=s}^k \sigma_j^2(A - \hat{C}_{:,i}^{(\text{Alg})} \hat{A}_{i,:}^{(\text{Alg})}), \quad i = 1, 2, 3, \quad (\text{B.4})$$

$$\varepsilon = \sum_{j=1}^k \sigma_j^2(A - \hat{C}^{(\text{Alg})} \hat{A}^{(\text{Alg})}) \quad (\text{B.5})$$

mit $\sigma_i(B)$ dem i -ten Singulärwert einer Matrix B zwei Möglichkeiten die Ergebnisse untereinander zu vergleichen. Dabei sind weiterhin

$$\left(\sum_{j=i}^k \sigma_j^2(A) \right)_{i=1,\dots,4} = \begin{pmatrix} 8.5151 \text{ E} + 00 \\ 7.8534 \text{ E} - 01 \\ 2.5372 \text{ E} - 02 \\ 3.3835 \text{ E} - 03 \end{pmatrix}$$

Vergleichsgrößen für Tabelle B.4.

Von den Werten ε_i aus (B.4) und ε aus (B.5) wird jeweils eine Größenordnung im Bereich der Störungen erwartet, da sich bei exakten Daten und einer exakten Bestimmung des Faktorenpaares $(\hat{c}_i(t), \hat{a}_i(x))$ beziehungsweise $(\hat{C}_{:,i}, \hat{A}_{i,:})$ eine Matrix $B = A - \hat{C}_{:,i} \hat{A}_{i,:}$ mit $\text{rank}(B) = s - 1$ ergeben würde.

	NNMF ¹	SPEXFA ²	BTEM ³	PCD
ε_1	3.7421 E−02	3.3835 E−03	(5.8107 E−03)	6.8964 E−03
ε_2	2.8500 E−02	3.3835 E−03	(5.8887 E−03)	6.4321 E−03
ε_3	1.5484 E−01	3.3835 E−03	(1.2617 E−02)	7.2720 E−03
ε	2.0345 E+00	3.3835 E−03	(3.9492 E−01)	1.2998 E−02

Tabelle B.4: Die Vergleichsgrößen aus (B.4) und (B.5).

B.3 Die Bildung von Hafnazyklopentan

Die Ergebnisse aus den Veröffentlichungen „*Kinetic and mechanistic investigations in homogeneous catalysis using operando UV/Vis spectroscopy*“ [20] und „*Combination of spectroscopic methods: In situ NMR and UV/Vis measurements to understand the formation of group 4 metallacyclopentanes from the corresponding metallacyclopropenes*“ [4] sind Gegenstand der Dissertation von Christian Fischer. Dabei sind in [20] Ergebnisse der Methode PCD dargelegt. In diesem Abschnitt werden einige der beschriebenen Methoden zur Bestimmung einer regularisierten Niedrigrangapproximation auf den zugehörigen Datensatz angewendet und die Resultate verglichen.

B.3.1 Systemfakten und Mehrdeutigkeit der Faktorisierung

Zu dem System selbst und der Eindeutigkeitsfrage bezüglich der Faktorisierung dessen Datenmatrix ist Folgendes bekannt.

¹Niedrigrangapproximation berechnet mit der MATLAB Routine `nmfans_comb` von Kim und Park.

²Niedrigrangapproximation berechnet mit der MATLAB Routine `SPEXFA` aus der begleitenden CD von [63].

³Zur Berechnung wurde $\hat{C}_i^{(\text{BTEM})}$ extern über eine lineare Ausgleichsrechnung bestimmt.

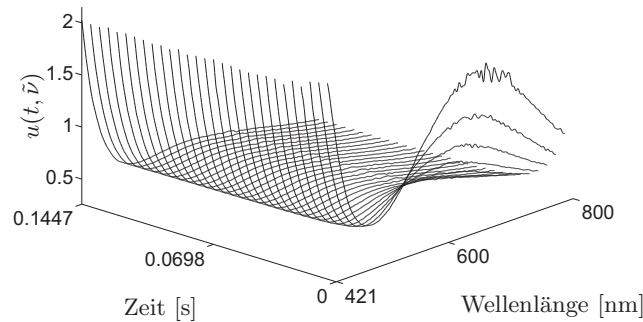


Abbildung B.7: Systemausschnitt, dargestellt sind nur die ersten 30 Spektren, da sich hierauf der Hauptteil der mit dem bloßen Auge erkennbaren Dynamik konzentriert.

Systemfakten:

- PD Dr. D. Heller, Dr. C. Fischer: Bildung von Hafnazyklopentan,
- Veröffentlicht in *Catalysis Today*: „Kinetic and mechanistic investigations in homogeneous catalysis using operando UV/Vis spectroscopy“ [20],
- Reaktionstemperatur: 25° C, Anfangskonzentration Edukt: 0.013 mol l⁻¹,
- Spektroskopieart: UV/Vis,
- Betrachteter Ausschnitt: $t \in [0 \text{ ms}, 2.49 \text{ s}]$, $\tilde{\nu} \in [420 \text{ nm}, 800 \text{ nm}]$,
- Dimensionen: $k = 500$, $n = 381$, rekonstruierbare Komponenten im Spektralbereich: $s = 3$,
- Komponenten: rac-(ebthi)Hf(C₄H₈), rac-(ebthi)Hf(η^2 -Me₃SiC₂SiMe₃), Intermediat,
- Abbildung B.7.

Aussagen bezüglich der Mehrdeutigkeit der Faktorisierung:

- Es liegt keine Eindeutigkeit der Niedrigrangapproximation vor.
- Systemrestriktion: $\hat{c}_1(\bar{t}_1) = 0.013$, $\hat{c}_2(\bar{t}_1) = \hat{c}_3(\bar{t}_1) = 0$.
- Annahme: zusätzlich gilt $\hat{c}_1(\bar{t}_{500}) = \hat{c}_2(\bar{t}_{500}) = 0$, $\hat{c}_3(\bar{t}_{500}) = 0.013$, das heißt es gilt unter Vernachlässigung der Skalierung damit $\hat{A}_{1,:} = A_{1,:}$, $\hat{A}_{3,:} = A_{500,:}$.
- Mit dieser Annahme lässt sich Satz 3.19 anwenden, das heißt $\hat{C}_{:,2}$ ist bis auf die Skalierung bekannt.

B.3.2 Lösungseinschließung durch Systemrestriktionen

Aus der Systemrestriktion nur eines Edukts und der Annahme nur eines Produkts ergibt sich ohne Rücksicht auf Sortierung und Skalierung $\hat{C}_{1,:} = (\alpha_1, 0, 0)$, $\hat{C}_{500,:} = (0, 0, \alpha_2)$ mit $\alpha_1, \alpha_2 > 0$. Damit sind gleichzeitig unskaliert $\hat{A}_{1,:}$ und $\hat{A}_{3,:}$ bekannt sowie $\hat{C}_{:,2}$.

Aus der Menge aller für das System relevanten Faktorisierungen ergeben sich mit Hilfe von Korollar 3.28 sowie den Sätzen 3.21 und 3.25 eindimensionale Kontinua für die restlichen Teilfaktoren. In Abbildung B.8 sind diese für $\hat{C}_{:,1}$ und $\hat{C}_{:,3}$ dargestellt.

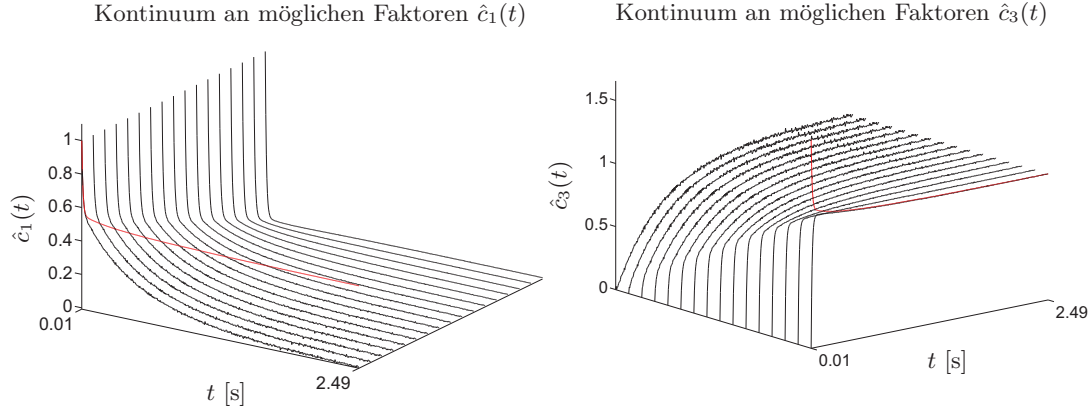


Abbildung B.8: Die möglichen Faktoren $\hat{c}_1(t)$ und $\hat{c}_3(t)$. Die rote Linie bildet jeweils eine obere Einhüllende, vergleiche auch Korollar 3.28.

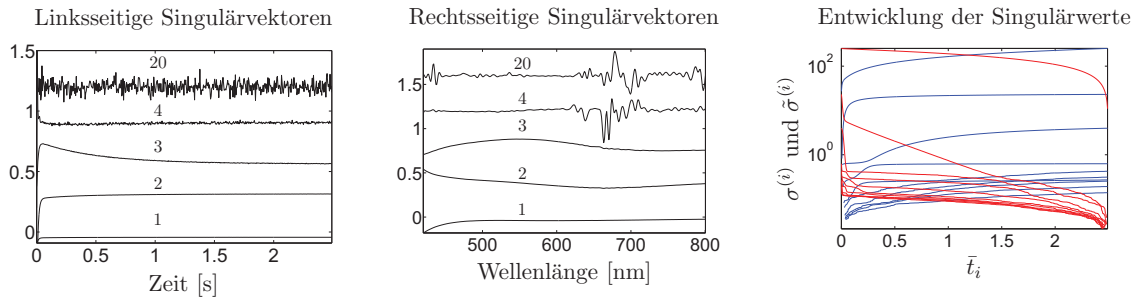


Abbildung B.9: Darstellung der links- und rechtsseitigen Singularvektoren aus der Singularwertzerlegung der Datenmatrix. Weiter ist rechts die Entwicklung der ersten 10 Singularwerte in Abhängigkeit von der zugrunde liegenden Datenbasis an Spektren dargestellt. Mit einer Betrachtung in positiver und in negativer Zeitrichtung ergeben sich $\sigma^{(i)}, \tilde{\sigma}^{(i)} \in \mathbb{R}^{10}$, $\sigma^{(i)} = \sigma_{1:10}(A_{1:i,:})$, $\tilde{\sigma}^{(i)} = \sigma_{1:10}(A_{i+1:k,:})$.

B.3.3 Analyse der Daten und Wahl der Parameter

Analog zu den beiden vorherigen Abschnitten erfolgt eine knappe Analyse des Datensatzes und basierend darauf eine Parameterwahl für die Anwendung der Algorithmen zur Bestimmung einer Niedrigrangapproximation.

i	σ_i	$\ A\ _F^{-2} \sum_{j=i+1}^k \sigma_j^2$	$\frac{1}{\ U_{:,i}\ _1} \sum_{j=1}^k U_{j,i}$	$\left\ \frac{d^2 U_{:,i}}{dt^2} \right\ _2$	$\frac{1}{\ V_{:,i}\ _1} \sum_{j=1}^n V_{j,i}$	$\left\ \frac{d^2 V_{:,i}}{d\bar{v}^2} \right\ _2$
1	2.501 E+02	5.403 E−03	−1.000 E+00	6.864 E−03	−1.000 E+00	8.105 E−04
2	1.801 E+01	2.437 E−04	1.762 E−01	1.204 E−01	−4.852 E−01	6.246 E−03
3	3.805 E+00	1.343 E−05	−7.758 E−03	1.257 E−01	1.091 E−01	1.191 E−02
4	4.257 E−01	1.054 E−05	−1.042 E−03	2.605 E+00	−2.283 E−01	2.816 E−01
5	3.147 E−01	8.970 E−06	−4.536 E−04	2.147 E+00	−1.595 E−01	3.444 E−01
6	3.126 E−01	7.416 E−06	−3.533 E−04	1.907 E+00	1.751 E−01	3.458 E−01
7	2.607 E−01	6.335 E−06	−1.341 E−03	1.672 E+00	2.225 E−01	2.244 E−01
8	1.925 E−01	5.745 E−06	−7.945 E−04	2.680 E+00	2.268 E−02	3.900 E−01
9	1.392 E−01	5.437 E−06	−6.063 E−04	2.560 E+00	1.821 E−03	3.514 E−01
10	1.305 E−01	5.166 E−06	8.296 E−05	2.587 E+00	−6.754 E−03	3.732 E−01
20	9.167 E−02	3.397 E−06	−2.461 E−05	2.561 E+00	−2.275 E−02	3.015 E−01

Tabelle B.5: Analyse des Datensatzes mittels Singularwertzerlegung.

Für die Bestimmung von Niedrigrangapproximationen des vorliegenden Datensatzes wird anhand von Tabelle B.5 und Abbildung B.9 der Parameter $s = 3$ sowie weiter für die Algorithmen

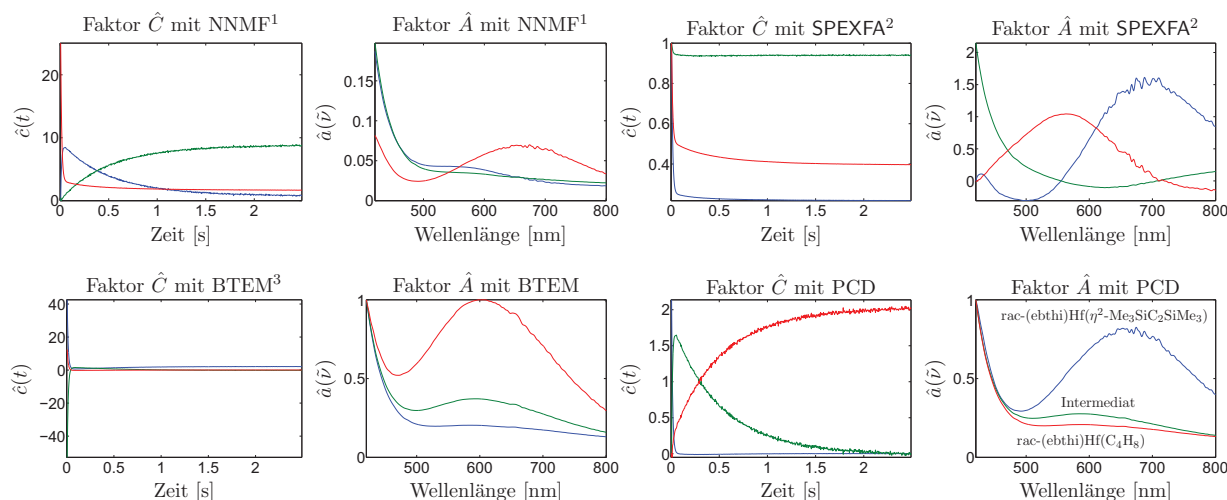


Abbildung B.10: Niedrigrangapproximationen des vorliegenden Datensatzes mittels verschiedener Algorithmen.

BTEM und PCD der Parameter $z = 3$ gewählt.

B.3.4 Ergebnisse der regularisierten Niedrigrangapproximation

Zur Lösung der Faktorisierungsaufgabe bezüglich der Datenmatrix werden eine klassische NNMF Methode sowie der SPEXFA-, der BTEM- und der PCD-Algorithmus angewendet. Die Ergebnisse sind in Abbildung B.10 dargestellt.

Anmerkungen zu den Ergebnissen und Abbildung B.10:

1. Der Parametersatz des BTEM-Algorithmus (Zielbereiche (ZB) aus Definition 5.6 und die Parameter aus (5.17)):
 - Linie 1 (blau): ZB: [420, 450]; $\delta_{Sh} = \delta_{Int} = 0$, $\delta_{Der} = 1$,
 - Linie 2 (grün): ZB: [500, 550]; $\delta_{Sh} = \delta_{Int} = 0$, $\delta_{Der} = 1$,
 - Linie 3 (rot): ZB: [600, 750]; $\delta_{Sh} = \delta_{Int} = 0$, $\delta_{Der} = 1$.
2. Der Parametersatz des PCD-Algorithmus in (6.11): $\gamma_1 = \gamma_2 = \gamma_3 = 1000$, $\gamma_7 = 1$, $\gamma_{11} = 5$, mit $I = \{1, 500\}$ und $c^{(s)} = (2, 2)^T$ aus (6.10). Weiter ist $\gamma_i = 0$ für $i \in \{4, 5, 6, 8, 9, 10, 12, 13\}$.
3. In dem Ergebnis von PCD weisen die Absorptivitäten $\hat{A}_{2,:}$ (grün) und $\hat{A}_{3,:}$ (rot) eine sehr hohe Stichproben-Kovarianz auf: $\text{Cov}(\hat{A}_{2,:}, \hat{A}_{3,:}) = 1.99 \text{ E} - 02$, bei $\text{Var}(\hat{A}_{2,:}) = 1.92 \text{ E} - 02$ und $\text{Var}(\hat{A}_{3,:}) = 2.10 \text{ E} - 02$.

B.3.5 Analyse der Ergebnisse

Da auch für dieses System keine eindeutige Äquivalenzklasse an Niedrigrangapproximationen existiert und ein Vergleich der Ergebnisse aus mathematischer Sicht nicht objektiv ist, werden erneut die Größen aus (B.4) und (B.5) herangezogen. Aus chemischer Sicht ist für das untersuchte System nur das Faktorisierungsergebnis von PCD relevant.

¹Berechnet mit der MATLAB Routine `nmfans_comb` von Kim und Park.

²Berechnet mit der MATLAB Routine `SPEXFA` aus der begleitenden CD von [63].

³Im Nachhinein durch lineare Ausgleichsrechnung auf den Zielbereichen bestimmt.

	NNMF ¹	SPEXFA ²	BTEM ³	PCD
ε_1	8.4450 E-01	8.4422 E-01	(8.5099 E-01)	8.4422 E-01
ε_2	8.4450 E-01	8.4422 E-01	(8.5071 E-01)	8.4422 E-01
ε_3	8.4422 E-01	8.4422 E-01	(8.5056 E-01)	8.4422 E-01
ε	8.4450 E-01	8.4422 E-01	(8.5942 E-01)	8.4422 E-01

Tabelle B.6: Auswertungen zu den Ergebnissen der verschiedenen Faktorisierungsmethoden. Die Bezeichnungen und Anmerkungen sind analog zu Tabelle B.4. Die Referenzwerte sind dazu

$$\left(\sum_{j=i}^k \sigma_j^2(A) \right)_{i=1,\dots,4} = \begin{pmatrix} 6.2875 \text{ E} + 04 \\ 3.3970 \text{ E} + 02 \\ 1.5324 \text{ E} + 01 \\ 8.4422 \text{ E} - 01 \end{pmatrix}.$$

B.4 Die Clusterbildung aus $\text{Rh}(\text{acac})(\text{CO})_2$

Im Zuge der Forschungen am Leibniz Institut für Katalyse in Rostock und zur Weiterentwicklung des Programmpakets PCD wurde ein Versuch zur Clusterbildung aus $\text{Rh}(\text{acac})(\text{CO})_2$ durchgeführt.

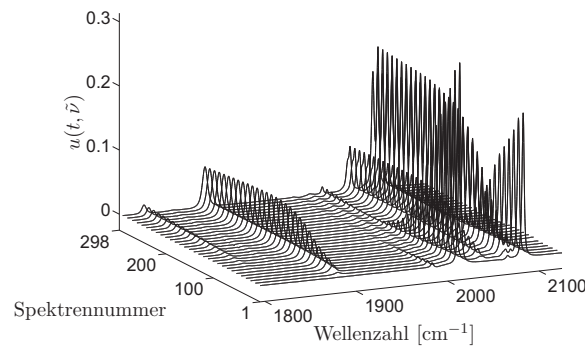


Abbildung B.11: Ausschnitt des Datensatzes aus Anhang B.4 (in der Abbildung ausgedünnt auf jedes zehnte Spektrum).

B.4.1 Systemfakten und Mehrdeutigkeit der Faktorisierung

Zu dem System selbst und der Eindeutigkeitsfrage bezüglich der Faktorisierung dessen Datenmatrix ist Folgendes bekannt.

Systemfakten:

- Dipl. Chem. E. Barsch: Clusterbildung $\text{Rh}(\text{acac})(\text{CO})_2$,
- Reaktionstemperatur: 30°C, Druck: 20 bar Synthesegas, Anfangskonzentration $\text{Rh}(\text{acac})(\text{CO})_2$: $0.66 \cdot 10^{-3} \text{ mol l}^{-1}$ Lösungsmittel: Cyclohexan,
- Spektroskopieart: FT-IR,
- Betrachteter Ausschnitt: $\tilde{\nu} \in [1795 \text{ cm}^{-1}, 2125 \text{ cm}^{-1}]$,
- Dimensionen: $k = 298$, $n = 2739$, rekonstruierbare Komponenten im betrachteten Spektralbereich: $s = 4$ (beziehungsweise $s = 3$ ohne Hintergrundspektrum),

¹Berechnet mit der MATLAB Routine `nmfals_comb` von Kim und Park.

²Berechnet mit der MATLAB Routine SPEXFA aus der begleitenden CD von [63].

³Zur Berechnung wurde $\hat{C}_i^{(\text{BTEM})}$ extern über eine lineare Ausgleichsrechnung bestimmt.

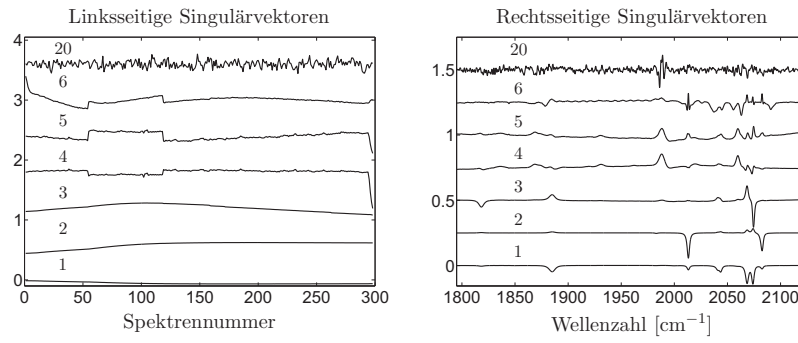


Abbildung B.12: Darstellung der links- und rechtsseitigen Singulärvektoren der Datenmatrix.

- Vorherrschende Komponenten im betrachteten Spektralbereich: $Rh_4(CO)_{12}$, $Rh_6(CO)_{16}$, $Rh(acac)(CO)_2$ sowie Lösungsmittel Cyclohexan,
- Es liegt nur ein Edukt ($Rh(acac)(CO)_2$) vor,
- Leichte Störung durch Subtraktion des Hintergrundspektrums: Es gibt i_0 mit $c_{i_0}(t) \not\equiv 0$, dies ließe sich über (6.7) einarbeiten,
- Die Absorptionen des Hintergrundspektrums (hauptsächlich Cyclohexan) treten nur in 4 der 298 Spektren mit einer absoluten Absorption von mehr als 1%, gemessen an der Maximalabsorption der Messreihe, auf,
- Abbildung B.11.

Aussagen bezüglich der Mehrdeutigkeit der Faktorisierung:

- Es liegt keine Eindeutigkeit der Niedrigrangapproximation vor.
- Abgesehen von Störungen gilt: $\hat{C}_{1,:} = (\alpha, 0, 0)$ mit $\alpha > 0$ sowie es gibt $i_1, i_2, i_3 \in \{1, \dots, n\}$ und $\varepsilon \in (\mathbb{R}_+ \setminus \{0\})^3$ mit $\hat{A}_{j,i_\ell} = \delta_{j,\ell\varepsilon_j}$ für $j, \ell = 1, 2, 3$. Das heißt Satz 3.15 ist anwendbar und $\hat{A}_{1,:}$ ist bis auf Störungen und Skalierung eindeutig in der Niedrigrangapproximation.

B.4.2 Analyse der Daten und Wahl der Parameter

Analog zu den drei vorherigen Abschnitten erfolgt eine knappe Analyse des Datensatzes und basierend darauf eine Parameterwahl für die Anwendung der Algorithmen zur Bestimmung einer Niedrigrangapproximation.

i	σ_i	$\ A\ _F^{-2} \sum_{j=i+1}^k \sigma_j^2$	$\frac{1}{\ U_{:,i}\ _1} \sum_{j=1}^k U_{j,i}$	$\left\ \frac{d^2 U_{:,i}}{dt^2} \right\ _2$	$\frac{1}{\ V_{:,i}\ _1} \sum_{j=1}^n V_{j,i}$	$\left\ \frac{d^2 V_{:,i}}{d\bar{\nu}^2} \right\ _2$
1	2.139 E+01	1.807 E-01	-1.000 E+00	2.239 E-03	-9.456 E-01	7.086 E-03
2	9.987 E+00	2.137 E-03	-4.064 E-01	3.093 E-03	-5.371 E-01	9.495 E-03
3	1.022 E+00	2.661 E-04	-5.460 E-03	2.037 E-02	-2.700 E-01	1.289 E-02
4	3.341 E-01	6.621 E-05	-9.897 E-03	3.541 E-01	6.908 E-01	6.574 E-03
5	1.826 E-01	6.509 E-06	4.852 E-03	4.387 E-01	-4.698 E-01	1.185 E-02
6	3.658 E-02	4.113 E-06	-1.480 E-03	2.827 E-01	-5.686 E-02	3.609 E-02
7	2.721 E-02	2.788 E-06	4.856 E-04	1.131 E+00	7.331 E-02	6.378 E-02
8	1.868 E-02	2.163 E-06	-4.344 E-04	4.428 E-01	1.235 E-01	5.009 E-02
9	1.185 E-02	1.912 E-06	-4.753 E-04	5.289 E-01	1.812 E-01	1.514 E-01
10	8.720 E-03	1.776 E-06	2.890 E-05	2.127 E+00	6.781 E-02	8.941 E-02
20	3.464 E-03	1.297 E-06	-2.057 E-05	2.154 E+00	-3.872 E-02	1.986 E-01

Tabelle B.7: Analyse des Datensatzes mittels Singulärwertzerlegung.

Für die Bestimmung von Niedrigrangapproximationen der vorliegenden Datenmatrix wird die Störung durch die ungenügende Lösung der Hintergrundproblematik vernachlässigt und anhand

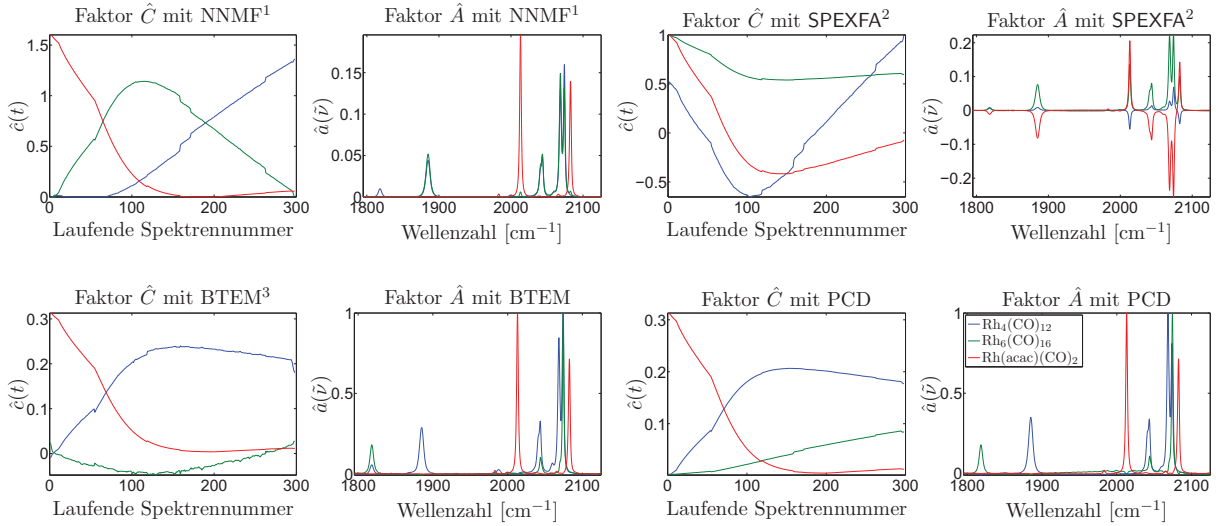


Abbildung B.13: Niedrigrangapproximationen der vorliegenden Datenmatrix.

von Tabelle B.7 und Abbildung B.12 der Parameter $s = 3$ sowie weiter für die Algorithmen BTEM und PCD der Parameter $z = 5$ gewählt. Auch $z = 4$ wäre vertretbar.

B.4.3 Die Ergebnisse der regularisierten Niedrigrangapproximation

Für den vorliegenden Datensatz werden die in den Kapiteln 5 und 6 beschriebenen Methoden eingesetzt, um Niedrigrangapproximationen der Datenmatrix zu bestimmen. Dargestellt sind die einzelnen Ergebnisse in Abbildung B.13.

Anmerkungen zu den Ergebnissen und Abbildung B.13:

1. Die durch den NNMF- und den BTEM-Algorithmus erzeugten Niedrigrangapproximationen erfüllen zwar die Bedingungen aus Faktorisierungsaufgabe 2.4, nach Rücksprache mit DC E. Barsch handelt es sich jedoch um keine chemisch korrekten Faktorisierungen.
2. Der Parametersatz des BTEM-Algorithmus (Zielbereiche (ZB) aus Definition 5.6 und die Parameter aus (5.17)):
 - Linie 1 (blau): ZB: [2038, 2046]; $\delta_{Sh} = \delta_{Int} = 0$, $\delta_{Der} = 1$,
 - Linie 2 (grün): ZB: [2080, 2085]; $\delta_{Sh} = \delta_{Int} = 0$, $\delta_{Der} = 1$,
 - Linie 3 (rot): ZB: [2005, 2022]; $\delta_{Sh} = 0$, $\delta_{Int} = \delta_{Der} = 1$.
3. Der Parametersatz des PCD-Algorithmus in (6.11): $\gamma_1 = \gamma_2 = 4$, $\gamma_3 = 1$, $\gamma_5 = 0.25$, $\gamma_6 = 0.01$, $\gamma_{12} = 0.1$. Weiter ist $\gamma_i = 0$ für $i \in \{4, 7, 8, 9, 10, 11, 13\}$.
4. In dem Faktorisierungsergebnis von PCD sind bei den (nicht eindeutigen) Absorptivitäten $\hat{A}_{1,\cdot}$ (blau) und $\hat{A}_{2,\cdot}$ (grün) starke Überlagerungen vorhanden: $\tilde{\nu} = 2043.5 \text{ cm}^{-1}$ und 2044.1 cm^{-1} sowie $\tilde{\nu} = 2073.7 \text{ cm}^{-1}$ und 2074.4 cm^{-1} .
5. Wie in den Vorbemerkungen erläutert, ist in einer nichtnegativen Matrixfaktorisierung des Datensatzes die Linie $A_{1,\cdot}$ eindeutig. In den Darstellungen aus Abbildung B.13 ist dies jeweils die rote Linie. Bei der Zerlegung des SPEXFA-Algorithmus sind einige Einträge in den Faktoren negativ, daher ist die besagte Linie nicht in der Faktorisierung enthalten.

¹Berechnet mit der MATLAB Routine `nmfanls_comb` von Kim und Park.

²Berechnet mit der MATLAB Routine `SPEXFA` aus der begleitenden CD von [63].

³Im Nachhinein durch lineare Ausgleichsrechnung auf den Zielbereichen bestimmt.

B.4.4 Die Analyse der Ergebnisse

Ebenfalls werden für eine Auswertung der Niedrigrangapproximationen die Größen aus (B.4) und (B.5) herangezogen.

	NNMF ¹	SPEXFA ²	BTEM ³	PCD
ε_1	1.6437 E−01	1.4863 E−01	(6.0923 E−01)	1.5905 E−01
ε_2	1.5773 E−01	1.4863 E−01	(2.4459 E−01)	1.4982 E−01
ε_3	1.5773 E−01	1.4863 E−01	(1.5176 E−01)	1.4982 E−01
ε	3.1352 E−01	1.4863 E−01	(1.6581 E+00)	1.6075 E−01

Tabelle B.8: Auswertungen zu den Ergebnissen der verschiedenen Methoden. Die Bezeichnungen sowie Anmerkungen sind analog zu Tabelle B.4. Ein Referenzwert ist $\sum_{i=4}^k \sigma_i^2(A) = 0.14863$.

B.5 Analyse der algorithmischen Ideen

In diesem Abschnitt erfolgt eine kurze Analyse der Vorgehensweise der in Abschnitt 6.4 beschriebenen Multi-Level-Strategie und der in 6.5 weiterentwickelten Methode der sukzessiven Faktorisierung durch Extrahieren einzelner Komponenten.

Ausgangspunkt für folgende Untersuchungen ist der Datensatz aus Anhang B.4. Dessen zugehörige Absorptionsmatrix wird mit dem PCD-Algorithmus durch ein Produkt aus zwei Faktorisierungsmatrizen vom Rang drei approximiert.

B.5.1 Die Anwendung der Multi-Level-Strategie

Die Multi-Level-Strategie aus Abschnitt 6.4 kommt mit der Startstufe $\ell = 3$ und der Ausdünnung $\tau_i = 0, \chi_i = i, i = 0, 1, 2, 3$ zur Anwendung, das heißt in Stufe $\mathcal{S}^{(3)}$ (siehe Definition 6.13) liegt eine Datenbasis von $k = 298$ Spektren à $n^{(3)} = 306$ Datenpunkte zugrunde.

Analyse der einzelnen Dünngitterfunktionen

Abbildung B.14 deutet zunächst die Schwierigkeit der Quadratmittelloptimierung von $F^{(l)}$ aus (6.27) an. Dazu sei $T^{(3,0)} \in \mathbb{R}^{3 \times 5}$ eine zufällig gewählte Matrix zur Initialisierung der Folge aus (6.28). Weiter seien (mit $e_i \in \mathbb{R}^s$ und $\bar{e}_i \in \mathbb{R}^z$ jeweils dem i -ten Einheitsvektor) folgende Funktionen $\varsigma_{i,j}^{(l)}, \zeta_{i,j}^{(l)} : \mathbb{R} \rightarrow \mathbb{R}$ definiert

$$\varsigma_{i,j}^{(l)}(x) = F^{(l)}(T^{(l,0)} + (x - T_{i,j}^{(l,0)})e_i\bar{e}_j^T), \quad i = 1, \dots, 3, \quad j = 1, \dots, 5, \quad l = 0, \dots, 3, \quad (\text{B.6})$$

$$\zeta_{i,j}^{(l)}(x) = F^{(l)}(\hat{T}^{(l)} + (x - \hat{T}_{i,j}^{(l)})e_i\bar{e}_j^T), \quad i = 1, \dots, 3, \quad j = 1, \dots, 5, \quad l = 0, \dots, 3, \quad (\text{B.7})$$

welche das Verhalten der Zielfunktion $F^{(l)}$ bezüglich der Variation jeweils eines Eintrags $T_{i,j}$ darstellen, betrachtet einerseits am Startpunkt der Iteration $T^{(l,0)}$ (Funktion ς aus (B.6)) und andererseits an der Minimalstelle $\hat{T}^{(l)}$ (Funktion ζ aus (B.7)).

¹Berechnet mit der MATLAB Routine `nmfanls_comb` von Kim und Park.

²Berechnet mit der MATLAB Routine `SPEXFA` aus der begleitenden CD von [63].

³Zur Berechnung wurde $\hat{C}_i^{(\text{BTEM})}$ extern über eine lineare Ausgleichsrechnung bestimmt.

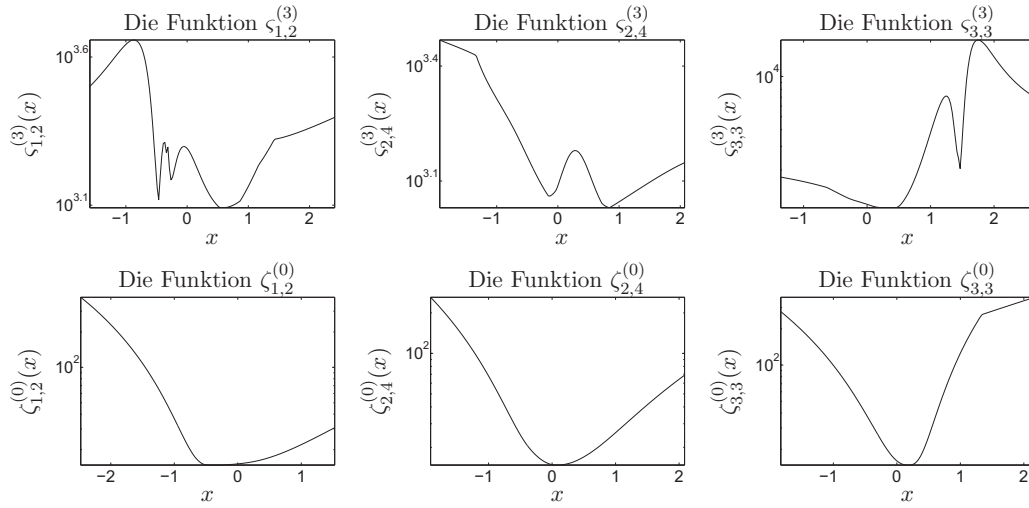


Abbildung B.14: Die Abbildung zeigt das Verhalten der Funktion $F^{(3)}$ im Startpunkt und der Funktion $F^{(0)}$ in der Minimalstelle, jeweils unter Variation einer Variablen anhand des Datensatzes zu Anhang B.4. Dargestellt sind einige der Funktionen $\zeta_{i,j}^{(3)}$ aus (B.6) für ausgewählte Einträge $T_{i,j}^{(3,0)}$ und eine zufällig gewählte Startmatrix $T^{(3,0)}$. Weiter ist als Vergleich dazu $\zeta_{i,j}^{(0)}$ aus (B.7) angegeben.

Analyse der Zusammenhänge unter den einzelnen Dünngitterfunktionen

Zur Untersuchung der Konvergenzraten in den einzelnen Grobgittern sei zunächst der Vektor $N \in \mathbb{N}^4$ mit N_l der Anzahl der in Stufe $\mathcal{S}^{(l)}$, $l = 0, \dots, 3$, benötigten Iterationen zur Quadratmittelminimierung definiert. Weiter seien mit der Nomenklatur aus (6.27) und (6.28) folgende Vektoren definiert

$$\theta_i^{(l)} = \left\| T^{(l,i)} - \hat{T}^{(l)} \right\|_F, \quad i = 0, \dots, N_l, \quad l = 0, \dots, 3, \quad (\text{B.8})$$

zur Analyse der Konvergenz der Iterierten $T^{(l,i)}$, $i = 1, \dots, N_l$, in den einzelnen Dünngitterfunktionen $F^{(l)}$,

$$\Theta_{i+\sum_{\nu=l+1}^3 N_\nu} = \left\| 2^{0.5(l-3)} T^{(l,i)} - \hat{T}^{(0)} \right\|_F, \quad i = 0, \dots, N_l, \quad l = 3, \dots, 0, \quad (\text{B.9})$$

zur Analyse der Konvergenz der Iterierten $T^{(l,i)}$, $i = 1, \dots, N_l$, $l = 3, \dots, 0$, unabhängig von den Dünngitterfunktionen $F^{(l)}$,

$$\omega_i^{(l)} = \left\| \nabla F^{(l)}(T^{(l,i)}) \right\|_F, \quad i = 0, \dots, N_l, \quad l = 0, \dots, 3, \quad (\text{B.10})$$

zur Analyse des Gradienten der einzelnen Dünngitterfunktionen $F^{(l)}$ und

$$\mu_i^{(l)} = \frac{\left\| T^{(l,i+1)} - T^{(l,i)} \right\|_F}{\left\| T^{(l,i)} \right\|_F}, \quad i = 0, \dots, N_l - 1, \quad l = 0, \dots, 3, \quad (\text{B.11})$$

zur Analyse des relativen Abstands zweier aufeinanderfolgender Iterierten eines Dünngitters $\mathcal{S}^{(l)}$. Eine grafische Darstellung der Einträge in diesen Vektoren liefert Abbildung B.15.

Die Prolongation

Für die Analyse der Prolongation zwischen zwei Dünngittern sind zwei Größen interessant. Der Abstand zwischen der Prolongation der Faktorisierung aus der größeren Stufe und der

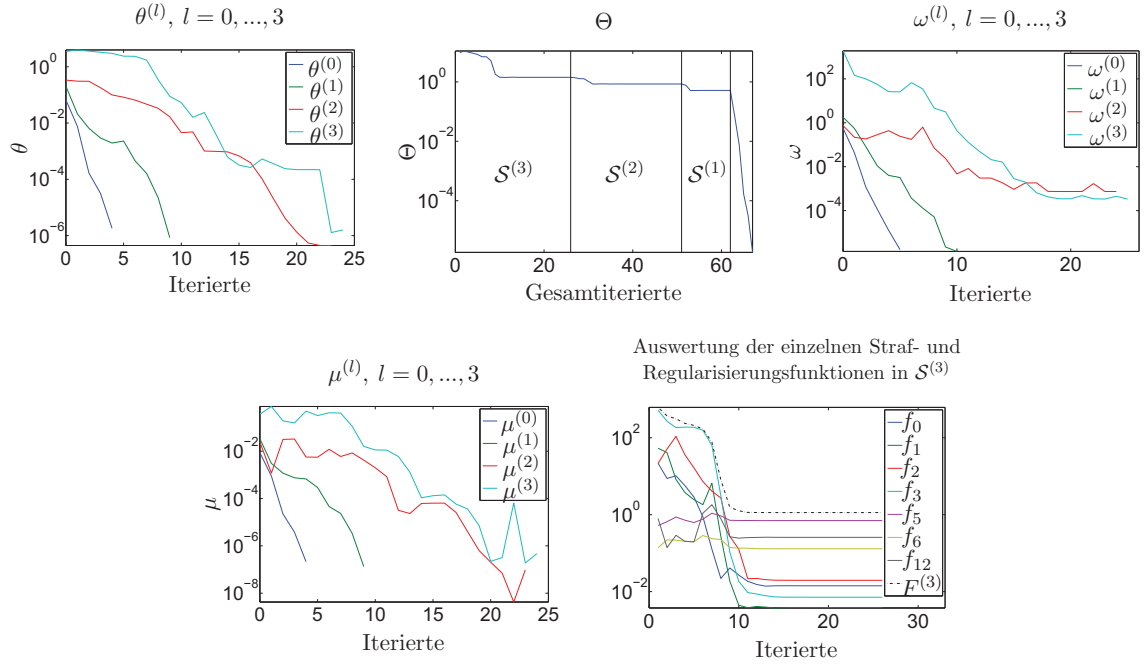


Abbildung B.15: Die Darstellung der Vektoren θ , Θ , ω und μ aus (B.8) bis (B.11) sowie der Betrag der einzelnen Straf- und Regularisierungsfunktionen aus (6.1) im Verlauf der Iteration in der Grobgitterstufe $\mathcal{S}^{(3)}$ rechts unten.

Faktorisierung aus der nächstfeineren lässt unter Berücksichtigung anderer Einflussfaktoren auf die Güte der Prolongation schließen. Dabei entspricht die Prolongation der Faktorisierung aus der gröberen Stufe dem Startvektor aus der feineren. Zudem zeigt ein Abstandsmaß für die Faktorisierungen aus zwei Grogittern, wie sinnvoll der Aufbau einer Dünngitterhierarchie für die Struktur eines konkreten Datensatzes ist.

Als Abstände zwischen den Prolongationen der jeweils gröberen Gitter und den Minimalstellen der feineren

$$\kappa_i = \|\rho_1(\hat{T}^{(i+1)}) - \hat{T}^{(i)}\|_F, \quad \tilde{\kappa}_i = \frac{\|\rho_1(\hat{T}^{(i+1)}) - \hat{T}^{(i)}\|_F}{\|\hat{T}^{(i)}\|_F}, \quad i = \ell - 1, \dots, 0,$$

ergeben sich in der untersuchten Anwendung

$$\begin{aligned} \kappa &= (6.7215 \text{ E-01}, 2.8874 \text{ E-01}, 6.6531 \text{ E-02})^T, \\ \tilde{\kappa} &= (7.2679 \text{ E-02}, 3.2265 \text{ E-02}, 7.8195 \text{ E-03})^T, \end{aligned}$$

wobei in diesem Fall stets ρ_1 der angewendete Prolongationsoperator ist. Die absoluten und relativen Differenzen der einzelnen mutmaßlichen Minimalstellen (skaliert wie in Bemerkung 6.20)

$$\begin{aligned} \Delta_i &= \|2^{0.5(i+1)}\hat{T}^{(i+1)} - 2^{0.5i}\hat{T}^{(i)}\|_F, \quad i = \ell - 1, \dots, 0, \\ \tilde{\Delta}_i &= \frac{\|2^{0.5(i+1)}\hat{T}^{(i+1)} - 2^{0.5i}\hat{T}^{(i)}\|_F}{\|2^{0.5i}\hat{T}^{(i)}\|_F}, \quad i = \ell - 1, \dots, 0, \end{aligned}$$

betragen

$$\begin{aligned} \Delta &= (6.9279 \text{ E-01}, 3.6115 \text{ E-01}, 5.1616 \text{ E-01})^T, \\ \tilde{\Delta} &= (7.4911 \text{ E-02}, 4.0357 \text{ E-02}, 6.0666 \text{ E-02})^T. \end{aligned}$$

Die Werte in κ , $\tilde{\kappa}$, Δ , $\tilde{\Delta}$ zeigen sehr deutlich, dass es für den behandelten Datensatz möglich und damit auch sinnvoll ist, eine Multi-Level-Strategie anzuwenden.

Analyse des Funktionsverhaltens in der letzten Iterierten

Die letzten Iterierten der Folgen aus (6.28) sind mutmaßliche Minimalstellen $\hat{T}^{(l)}$ für die einzelnen Dünngitterfunktionen $F^{(l)}$, $l = 3, \dots, 0$. Zur Analyse des Verhaltens der Funktionen $F^{(l)}$ in $\hat{T}^{(l)}$ ergeben sich die Einträge aus Tabelle B.9, welche im Rahmen numerischer Berechnungen jeweils klare Kennzeichen für lokale Minimalstellen sind.

l	$\ (\nabla F^{(l)})(\hat{T}^{(l)})\ _F$	$\ (\nabla F^{(l)})(\hat{T}^{(l)})\ _\infty$	$\lambda_{\max}(H_{F^{(l)}}(\hat{T}^{(l)}))$	$\lambda_{\min}(H_{F^{(l)}}(\hat{T}^{(l)}))$
3	3.3429 E-04	2.9860 E-04	1.0183 E+03	1.1791 E+00
2	7.3953 E-04	7.3050 E-04	2.4975 E+02	6.1094 E-01
1	1.3778 E-06	8.2358 E-07	9.0194 E+01	1.5798 E+00
0	1.6648 E-06	7.9775 E-07	1.0114 E+02	2.6283 E+00

Tabelle B.9: Die Werte wurden jeweils zu den letzten Iterierten $\hat{T}^{(l)}$ in den einzelnen Multi-Level-Stufen bezüglich des Datensatzes zu Anhang B.4 berechnet. Dabei gibt $\lambda_{\max}(B)$ den größten Eigenwert einer quadratischen Matrix B sowie $\lambda_{\min}(B)$ den kleinsten an und $H_{F^{(l)}}$ ist die Hessematrix der zweiten partiellen Ableitungen der Funktion $F^{(l)}$ nach den zu optimierenden Variablen. Für eine numerische nichtlineare Ausgleichsrechnung über 15 Variablen ist sowohl die notwendige als auch die hinreichende Bedingung für eine lokale Minimalstelle in der Wurzelstufe $S^{(0)}$ ausreichend erfüllt.

Vergleich der Ergebnisse für unterschiedliche Startgitter

In folgendem Beispiel soll die Multi-Level-Strategie für die zwei Versuchsreihen aus B.2 und B.4 unter Anwendung unterschiedlicher Startstufen untersucht werden. Das Ziel ist ein geeignetes Grobgitter $S^{(\ell)}$ derart auszuwählen, dass die Approximation zufriedenstellend stabil berechnet wird sowie gleichzeitig die Rechenzeit minimal ist.

Für die Anwendung wurde stets $l_t = 0$ gesetzt und l_x variiert. Für jeweils $m = 100$ zufällig erzeugte Startmatrizen $T^{(\ell,0)} \in \mathbb{R}^{s \times z}$ wurde die Anzahl der erzeugten Folgen von Iterierten gezählt, welche in eine mutmaßliche globale Minimalstelle der Startstufe führten. Daraus lässt sich statistisch auf die Wahrscheinlichkeit für eine Folge von Iterierten bei $m = 4$ oder $m = 10$ schließen, was durchaus praxisnah ist. Zusätzlich wurde die Rechenzeit für den gesamten Durchlauf der Grobgitterhierarchie gemessen. Die Werte der durchgeführten Rechnungen sind in den Tabellen B.10 und B.11 festgehalten.

Weiter lässt sich mit den Tabellen B.10 und B.11 eine Abhängigkeit darstellen, wie hoch die Gesamtrechenzeit ist im Vergleich zur statistischen Wahrscheinlichkeit dafür, bei $m = 4$ zufälligen Startmatrizen keine Folgen von Iterierten zu erzeugen, die in eine mutmaßliche globale Minimalstelle der Startstufe führen. Eine solche Darstellung findet sich in Abbildung B.16.

B.5.2 Die Anwendung der sukzessiven Niedrigrangapproximation

In Abschnitt 6.5 wurde eine Vorgehensweise der sukzessiven Faktorisierung basierend auf einer Folge von lokalen Niedrigrangapproximationen beschrieben. Dabei werden die s Faktorenpaare $(\hat{C}_{:,i}, \hat{A}_{i,:})$, $i = 1, \dots, s$, nacheinander einzeln extrahiert und abgespalten.

Grundlage dieser Vorgehensweise ist wieder eine Singulärwertzerlegung der Datenmatrix. Die Teilfaktoren $\hat{A}_{i,:}$ werden jeweils durch eine Linearkombination $T_i \in \mathbb{R}^{1 \times z}$ aus den ersten z rechtsseitigen Singulärvektoren $V_{:,1:z}$ bestimmt

$$\hat{A}_{i,:} = T_i V_{:,1:z}^T.$$

l_x	$\lfloor \frac{2455}{2^{l_x}} \rfloor$	Zufalls- startvek. m	# Iteration in globales Minimum	Stat. Wahrsch. auf mind. 1 erfolg. Iter. für $m = 10$	Stat. Wahrsch. auf mind. 1 erfolgr. Iter. für $m = 4$	Gesamt- rechenzeit $m = 100$
6	38	100	8	56.561 %	28.360 %	176.1 s
5	76	100	59	99.987 %	97.174 %	211.9 s
4	153	100	69	99.999 %	99.076 %	235.0 s
3	306	100	49	99.881 %	93.235 %	337.5 s
2	613	100	64	99.996 %	98.320 %	543.7 s
1	1227	100	74	> 99.999 %	99.543 %	944.9 s
0	2455	100	82	> 99.999 %	99.895 %	1886.6 s

Tabelle B.10: Die Auswertung betrifft die Versuchsreihe aus Anhang B.4 mit $s = 3$, $z = 5$. Eine Ausdünnung ist nur in x -Richtung durchgeführt, das heißt $\tau_i = 0$, $\chi_i = i$, $i = 0, \dots, l_x$. Mit $n = 2455$ Stützwerten für x in der Wurzelstufe ergibt sich damit eine Ausdünnung auf $n_{l_x} = \lfloor \frac{2455}{2^{l_x}} \rfloor$ in der jeweiligen Startstufe. Bei der erwähnten globalen Minimalstelle handelt es sich nur um eine Mutmaßung. Gerechnet wurde dabei auf einer Maschine mit einer PENTIUM 2.4 Ghz CPU und 4 GB Arbeitsspeicher.

l_x	$\lfloor \frac{292}{2^{l_x}} \rfloor$	Zufalls- startvek. m	# Iteration in globales Minimum	Stat. Wahrsch. auf mind. 1 erfolg. Iter. für $m = 10$	Stat. Wahrsch. auf mind. 1 erfolgr. Iter. für $m = 4$	Gesamt- rechenzeit $m = 100$
4	18	100	35	98.654 %	82.149 %	174.7 s
3	36	100	53	99.947 %	95.120 %	206.8 s
2	72	100	57	99.978 %	96.581 %	261.9 s
1	146	100	59	99.987 %	97.174 %	366.7 s
0	292	100	62	99.994 %	97.915 %	578.9 s

Tabelle B.11: Die Auswertung betrifft die Versuchsreihe aus Anhang B.2 mit $s = 3$, $z = 5$ und der Ausdünnung nur in x -Richtung, das heißt $\tau_0 = \dots = \tau_\ell = l_t = 0$.

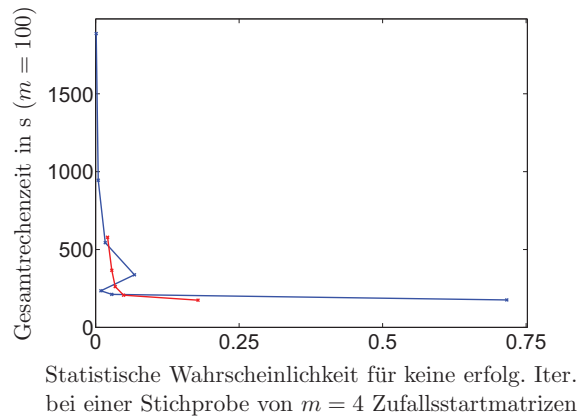


Abbildung B.16: Ausgangspunkt sind die Werte aus den Tabellen B.10 und B.11. In Abhängigkeit der Startstufe ist die Rechenzeit gegen die statistische Wahrscheinlichkeit auf einen Misserfolg bei einer Wahl von $m = 4$ zufälligen Startmatrizen für die Optimierung dargestellt, einerseits für $\mathcal{S}^{(6)}$ bis $\mathcal{S}^{(0)}$ (blaue Linie, Tabelle B.10) und andererseits für $\mathcal{S}^{(4)}$ bis $\mathcal{S}^{(0)}$ (rote Linie, Tabelle B.11). Zu erkennen ist dabei eine L-Kurve mit jeweils dem „Knick“ für $\mathcal{S}^{(5)}$ beziehungsweise $\mathcal{S}^{(3)}$, demzufolge können für diese Beispiele die Werte $l_x = 5$ beziehungsweise $l_x = 3$ als geeignet angesehen werden.

Der jeweils zugehörige Teilfaktor $\hat{C}_{:,i}$ wird als Linearkombination $S_i \in \mathbb{R}^z$ aus den ersten z linksseitigen Singulärvektoren U mit der zusätzlichen Gewichtung der ersten z Singulärwerte berechnet

$$\hat{C}_{:,i} = U \Sigma_{:,1:s} S_i,$$

vergleiche (6.31). Dabei lassen sich die Vektorenpaare (T_i, S_i) jeweils als eigenständige Variablen betrachten oder in der Abhängigkeit $S_i = S_i(T_i)$ aus (6.33).

Insgesamt ergibt die Zusammensetzung von $\tilde{T} = (T_1^T, \dots, T_s^T)^T$ und $\tilde{S} = (S_1, \dots, S_s)$ die beiden Transformationsmatrizen $\tilde{T} \in \mathbb{R}^{s \times z}$ und $\tilde{S} \in \mathbb{R}^{z \times s}$ aus (4.5) und damit die Faktorisierungsmatrizen

$$\begin{aligned} \hat{A} &= \tilde{T} V_{:,1:z}^T, \\ \hat{C} &= U \Sigma_{:,1:z} \tilde{S}. \end{aligned}$$

Ergebnisvergleich der gesamtheitlichen und der sukzessiven Faktorisierung

Angewendet auf den Datensatz aus Anhang B.4 ergibt sich eine nahezu identische Niedrigrangapproximation der Datenmatrix im Vergleich zu dem Ergebnis der Methode PCD aus Abschnitt B.4. Dazu sind in Abbildung B.17 die Differenzen

$$\hat{C}^{(\text{PCD})} - \hat{C}^{(\text{suk})}, \tag{B.12}$$

$$\hat{A}^{(\text{PCD})} - \hat{A}^{(\text{suk})} \tag{B.13}$$

der Faktoren aus der gesamtheitlichen ($\hat{C}^{(\text{PCD})}$, $\hat{A}^{(\text{PCD})}$, siehe Abbildung B.13) und der sukzessiven Faktorisierung ($\hat{C}^{(\text{suk})}$, $\hat{A}^{(\text{suk})}$) dargestellt.

Die Unterschiede in den Transformationen \hat{T} , \hat{T}^+ mit $\hat{T} = \hat{T}^{(\text{PCD})}$ aus dem Ergebnis zu Abbildung B.13 und \tilde{T} , \tilde{S} aus der Anwendung der sukzessiven Niedrigrangapproximation sind

$$\begin{aligned} \|\hat{T} - \tilde{T}\|_F^2 &= 2.42489 \text{ E} - 01, \\ \|\hat{T}^+ - \tilde{S}\|_F^2 &= 7.63402 \text{ E} - 03, \quad \|\Sigma_{:,1:5}(\hat{T}^+ - \tilde{S})\|_F^2 = 3.86585 \text{ E} - 02. \end{aligned}$$

Weiter sind folgende Kenngrößen zu \tilde{T} und \tilde{S} interessant

$$\begin{aligned} \|\tilde{T}\tilde{S} - I_3\|_F^2 &= 1.5796 \text{ E} - 05, \quad \|\tilde{S}\tilde{T} - \text{diag}(1, 1, 1, 0, 0)\|_F^2 = 1.6191 \text{ E} - 01, \\ \text{diag}(\tilde{S}\tilde{T}) &= (1.0025, 1.0001, 0.9435, 0.0198, 0.0393)^T, \\ \|\tilde{S}\tilde{T} - \text{diag}((\tilde{S}\tilde{T})_{1,1}, \dots, (\tilde{S}\tilde{T})_{5,5})\|_F^2 &= 1.5677 \text{ E} - 01, \\ \|\tilde{S} - \tilde{T}^+\|_F^2 &= 2.9325 \text{ E} - 03. \end{aligned}$$

Damit lässt sich eine Anwendbarkeit dieser Strategie auf den Datensatz B.4 begründen. Dazu sei angemerkt, dass die Datenstruktur der dünnbesetzten Faktoren sowohl in \hat{C} als auch in \hat{A} dies begünstigt.

B.6 Numerische Ergebnisse für eine Systemerweiterung

In den Abschnitten 4.4 und 4.5 wurde untersucht, unter welchen Voraussetzungen einige der Teilfaktoren eindeutig sind. Weiter wurde untersucht, welche indirekten Auswirkungen dies auf andere Faktoren hat. Sind Teilfaktoren für eine vorliegende Datenmatrix bekannt, so schränkt dies die Menge aller Faktorisierungen auf eine Menge der relevanten ein. Unter Umständen

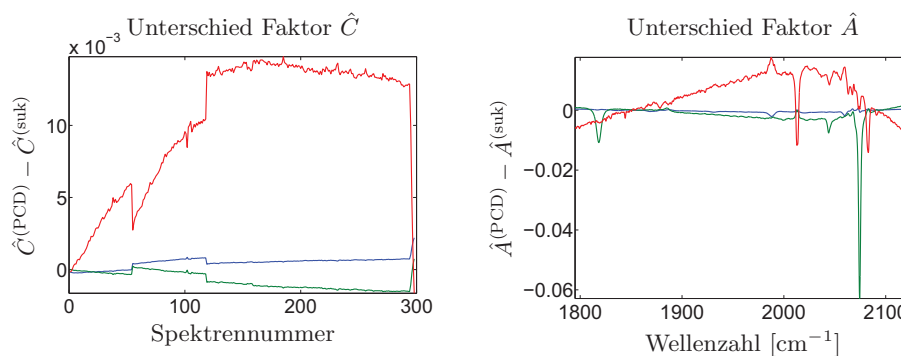


Abbildung B.17: Dargestellt ist der absolute Unterschied der Niedrigrangapproximationen aus (B.12) und (B.13). Zum Vergleich ist auf $\hat{C}^{(\text{PCD})}$ und $\hat{A}^{(\text{PCD})}$ in Abbildung B.13 verwiesen.

sind dadurch auch andere Teilfaktoren in Bezug auf Relevanz eindeutig. In den untersuchten Datensätzen sind unter Berücksichtigung von Störungen folgende Teilfaktoren eindeutig beziehungsweise durch Kenntnisse anderer Teilfaktoren eindeutig im Sinne von relevanten Faktorisierungen.

- In dem Datensatz zu B.1 sind die Faktoren $\hat{C}_{:,2}$ und $\hat{A}_{1,:}$ eindeutig (Sätze 3.15 und 3.19).
- In dem Datensatz zu B.3 ist durch die gemachten Annahmen ($\hat{A}_{1,:}$ und $\hat{A}_{3,:}$ bekannt) der Faktor $\hat{C}_{:,2}$ eindeutig (Satz 3.19).
- In dem Datensatz zu B.4 ist der Faktor $\hat{A}_{3,:}$ (rote Linie Bild NNMF und PCD) eindeutig (Satz 3.15).

B.6.1 Lösungseinschließung mittels Systemerweiterung

Weiter wurde in Korollar 4.24 ein Ansatz dargestellt, in dem an einen Datensatz mehrere einzelne bekannte Teilfaktoren angefügt werden.

Um dieses Vorgehen zu verdeutlichen, sei ein weiterer Datensatz aus der Spektroskopie angeführt. Zur Einschränkung der Lösungsvielfalt werden dabei die Absorptionsspektren \hat{A} von fünf Komponenten an die Messmatrix angefügt. Damit ist jeweils ein Teil der Voraussetzungen von Satz 3.15 pro Komponente erfüllt. Da die Teilfaktoren selbst den zweiten Teil der Voraussetzungen nicht erfüllen, ergibt sich keine endgültige Eindeutigkeit in der Niedrigrangapproximation, sondern nur eine Einschränkung an Auflösungsmöglichkeiten.

Präformierung eines hydroformylierungsaktiven Katalysators aus $\text{Rh}(\text{acac})(\text{CO})_2$

Als Anwendungsbeispiel sei ein Datensatz aus 6 verschiedenen Versuchabläufen von Dipl. Chem. E. Barsch gewählt.

Systemfakten und Mehrdeutigkeit der Faktorisierung

Zu dem System selbst und der Eindeutigkeitsfrage bezüglich der Faktorisierung dessen Datenmatrix ist Folgendes bekannt.

Systemfakten:

- Versuchsdurchführung und Datenaufnahme: Dipl. Chem. E. Barsch,
- Insgesamt Daten aus 6 verschiedenen Versuchabläufen zusammengesetzt,
- Spektroskopieart: FT-IR,

- Betrachteter Ausschnitt: $\tilde{\nu} \in [1795 \text{ cm}^{-1}, 2125 \text{ cm}^{-1}]$,
- Dimensionen: $k = 496$, $n = 2455$, rekonstruierbare Komponenten im betrachteten Spektralbereich: $s = 6$ (beziehungsweise $s = 5$ ohne Hintergrundspektrum),
- Leichte Störung durch Subtraktion des Hintergrundspektrums: Es gibt i_0 mit $c_{i_0}(t) \not\equiv 0$.

Aussagen bezüglich der Mehrdeutigkeit der Faktorisierung und Besonderheiten:

- Es liegt keine Eindeutigkeit der Niedrigrangapproximation vor.
- Starke Korrelation zweier Teilfaktoren $\hat{C}_{:,i}$ und $\hat{C}_{:,j}$.

Systemerweiterung

An die Datenmatrix $A \in \mathbb{R}^{496 \times 2455}$ werden zunächst fünf Teilfaktoren $\hat{A} \in \mathbb{R}^{5 \times 2455}$ angefügt

$$A' = \begin{pmatrix} A \\ \hat{A} \end{pmatrix} \quad (\text{B.14})$$

und eine Niedrigrangapproximation mit einem Produkt aus zwei Matrizen vom Rang 5 mittels des NNMF-Algorithmus von Kim und Park bestimmt

$$A' = \begin{pmatrix} \hat{C}' \\ M \end{pmatrix} \hat{A}', \quad \hat{C}' \in \mathbb{R}^{496 \times 5}, \quad \hat{A}' \in \mathbb{R}^{5 \times 2455}, \quad M \in \mathbb{R}^{5 \times 5}. \quad (\text{B.15})$$

Die negativen Einträge durch Störungen aufgrund der ungenügenden Subtraktion des Hintergrundspektrums werden durch die NNMF-Methode vernachlässigt und auf Null gesetzt, das sechste Faktorenpaar wird demzufolge nicht bestimmt. Durch die Systemrestriktion und die Art der Faktoren selbst ergibt sich eine Einschränkung der Menge an Faktorisierungen.

Weiter wird als Vergleich eine Niedrigrangapproximation der Datenmatrix A ohne Systemrestriktionen in fünf Faktoren bestimmt. Die Abbildung B.18 zeigt die Ergebnisse beider Berechnungen und damit die einschränkende Wirkung durch externe Systeminformationen und -restriktionen.

Zu der Faktorisierung von A' in ein Produkt wie in (B.15) ergibt sich

$$M = \begin{pmatrix} 1.007 & 0 & 0 & 0 & 0 \\ 0.026 & 0.997 & 0 & 0.003 & 0 \\ 0 & 0 & 0 & 0.999 & 0 \\ 0.002 & 0.086 & 0 & 0.092 & 0.988 \\ 0 & 0 & 0.996 & 0 & 0 \end{pmatrix}.$$

Die Abweichung der Matrix M von einer permutierten Einheitsmatrix analog zu (4.7) aus Bemerkung 4.26 ist gering. Als Abweichungen von den zu erwartenden Singulärwerten eins ergeben sich

$$(1 - \sigma_i(M))_{i=1,\dots,5} = (-0.0595, -0.0120, 0.0043, 0.0093, 0.0678).$$

Die Anwendung von PCD unter einer Systemerweiterung

Analog zu diesem Vorgehen lassen sich für die Berechnung einer Niedrigrangapproximation mit der Routine PCD sechs Teilfaktoren $\hat{A} \in \mathbb{R}^{6 \times 2455}$ an die Datenmatrix anfügen

$$A' = \begin{pmatrix} A \\ \hat{A} \end{pmatrix}. \quad (\text{B.16})$$

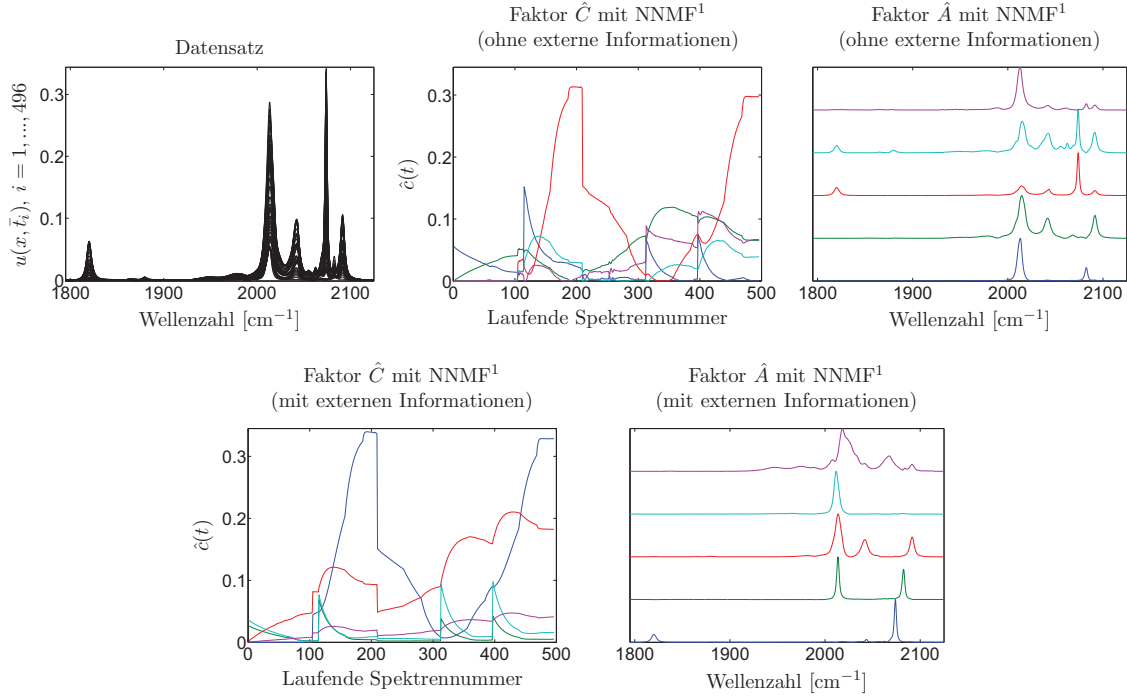


Abbildung B.18: Dargestellt sind zwei Niedrigrangapproximationen mit dem NNMF-Algorithmus von Kim und Park. Oben: die Faktorisierung der Matrix A mittels \hat{C} und \hat{A} , unten: Die Faktorisierung der Matrix A' aus (B.14) mit \hat{C}' und \hat{A}' aus (B.15).

Im Unterschied zu der Faktorisierung des Algorithmus von Kim und Park ergibt sich hierbei zusätzlich das Faktorenpaar des unvollständig beziehungsweise zu stark abgezogenen Hintergrundspektrums.

Die Ergebnisse der Niedrigrangapproximation unter einer Systemerweiterung

Die Wahl der Parameter zur Niedrigrangapproximation ist $s = 6$, $z = 7$ sowie zur Restriktion $\gamma_1 = \gamma_2 = 1000$, $\gamma_3 = 250$, $\gamma_5 = 1$ und $\gamma_i = 0$, $i \in \{4, 6, \dots, 13\}$. Erneut wird eine Multi-Level-Strategie mit $\ell = 3$ und der Ausdünnung $\tau_i = 0$, $\chi_i = i$, $i = 0, 1, 2, 3$, angewendet. Abbildung B.19 zeigt die Ergebnisse der Niedrigrangapproximation.

Als veränderter Teilfaktor \hat{C}' ergibt sich

$$A' = \begin{pmatrix} \hat{C}' \\ M \end{pmatrix} \hat{A}', \quad M = \begin{pmatrix} 1.0009 & 0.0251 & -0.0100 & 0.0182 & -0.0100 & -0.0087 \\ -0.0100 & 1.0000 & -0.0100 & -0.0100 & 0.0126 & -0.0100 \\ 0.0056 & 0.0029 & 1.0000 & -0.0100 & -0.0008 & -0.0100 \\ -0.0100 & 0.0043 & -0.0100 & 1.0000 & 0.0414 & -0.0094 \\ 0.0157 & -0.0100 & 0.0146 & -0.0105 & 1.0000 & -0.0101 \\ 0.0810 & -0.0065 & -0.0099 & -0.0100 & -0.0100 & 0.9967 \end{pmatrix}.$$

Die negativen Einträge in M sind dabei auf die in Straffunktion \tilde{f}_3 aus (6.7) umgesetzten Restriktionen $\hat{C}_{i,j} \geq c_{\min} = -0.01$ zurückzuführen. Die Abweichungen von den zu erwartenden Singulärwerten eins sind

$$(1 - \sigma_i(M))_{i=1, \dots, 6} = (-0.0397, -0.0181, -0.0052, -0.0034, 0.0211, 0.0441).$$

¹Berechnet mit der MATLAB Routine `nmfanls_comb` von Kim und Park.

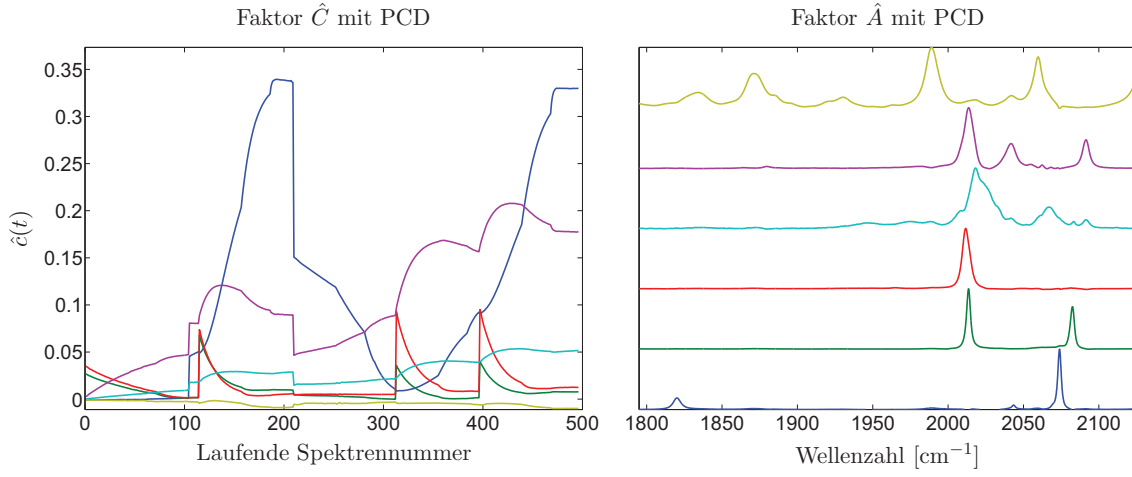


Abbildung B.19: Niedrigrangapproximation der Matrix A' aus (B.16) mittels PCD.

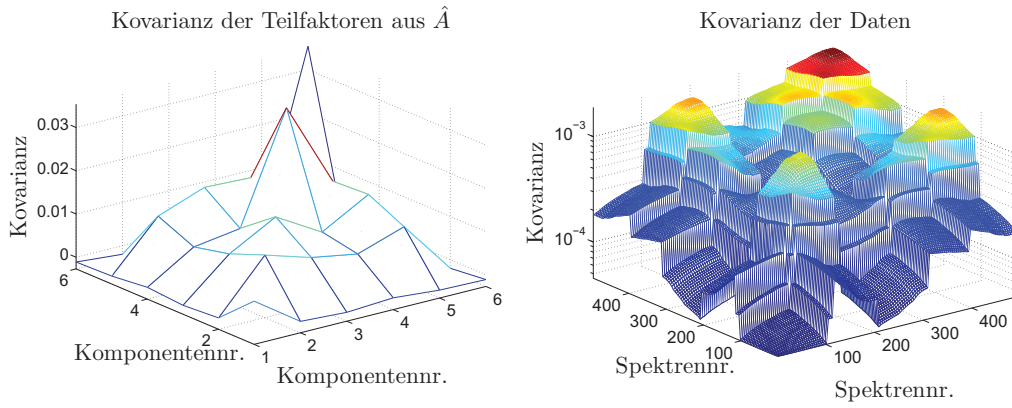


Abbildung B.20: Dargestellt sind die Stichproben-Kovarianzen $\text{Cov}(\hat{A}_{i,:}, \hat{A}_{j,:})$, $i, j = 1, \dots, 6$, der einzelnen extrahierten Teilfaktoren sowie die Stichproben-Kovarianzen der einzelnen Messungen $\text{Cov}(A_{i,:}, A_{j,:})$, $i, j = 1, \dots, 496$. Letztere ergeben sich dabei alle als positiv und sind daher mit logarithmischer Skala angegeben. Entscheidend sind nicht die absoluten Werte, sondern die relativen jeweils bezüglich den Diagonaleinträgen.

Die Korrelation der Daten

Stark korrelierte Daten erhöhen im Normalfall die Vielfalt aller möglichen Niedrigrangapproximation. Wie stark die Daten der in diesem Abschnitt vorliegenden Extinktionsmatrix korreliert sind, zeigen die Darstellungen aus Abbildung B.20. Dabei sind zum einen die Stichproben-Kovarianzen der extrahierten Teilfaktoren \hat{A} (Reinkomponentenspektren)

$$\text{Cov}(\hat{A}_{i,:}, \hat{A}_{j,:}), \quad i, j = 1, \dots, 6,$$

sowie die Stichproben-Kovarianzen der einzelnen Messungen

$$\text{Cov}(A_{i,:}, A_{j,:}), \quad i, j = 1, \dots, 496,$$

bestimmt worden.

Bemerkung B.1. Entscheidend in Abbildung B.20 rechts sind dabei nicht die Absolutwerte, sondern die Abweichungen von der Diagonalen.

B.7 Kritisches Fazit

Bei den betrachteten Datensätzen liegen stets Matrizen vor, welche keine eindeutige nichtnegative Faktorisierung besitzen. Aus mathematischer Sicht kann es demnach in der Faktorisierung keine Bewertung „richtig“ oder „falsch“ geben. Eine chemische Interpretation der Ergebnisse ist jedoch zulässig.

Diesbezüglich schneiden die Programmpakete BTEM und PCD bei der Anwendung auf Daten aus der FT-IR Spektroskopie am besten ab, auch weil die eingesetzten Restriktionen teilweise speziell auf die Faktorisierung von Daten dieser Spektroskopietechnik zugeschnitten sind und die Datenstruktur eine lokale Niedrigrangapproximation wie beim BTEM-Algorithmus zulässt.

Für die Anwendung in der UV/Vis Spektroskopie ergeben sich in der Regel erst durch die Einbindung externer Systemkenntnisse und damit verbundenen Restriktionen chemisch relevante Ergebnisse.

C Numerische Ergebnisse zur Systemanpassung

In Abschnitt 6.2 und Kapitel 7 wurde eine Aufgabenstellung der Chemometrie basierend auf dem Ergebnis einer regularisierten nichtnegativen Niedrigrangapproximation untersucht. Ausgangspunkt für die Daten war dabei die Beobachtung eines Reaktionsverlaufs.

Bei derartigen Daten stehen die Teilfaktoren der Konzentrationsprofile über eine gewöhnliche Differentialgleichung in Zusammenhang, welche unbekannte Parameter enthält. Diese Parameter sind von großem Interesse. In der gewonnenen nichtnegativen Matrixfaktorisierung sind die Einträge des Faktors \hat{C} Konzentrationen zu bestimmten Zeitpunkten und lassen sich damit zur Bestimmung der Parameter der Differentialgleichung nutzen.

C.1 Die Anpassung eines kinetischen Modells zum System aus Anhang B.2

Die Einträge des Faktors \hat{C} sind diskrete Auswertungen einer Funktion $c(t)$ der Konzentrationsprofile. Diese ist Lösung eines Anfangswertproblems mit einer Differentialgleichung

$$\dot{c}(t) = f(t, c, \phi),$$

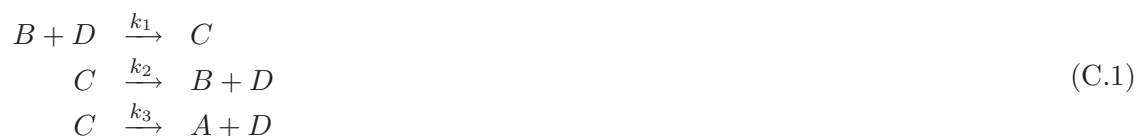
welche den Parameter $\phi \in \mathbb{R}^p$ enthält. Das Ziel ist die Anpassung des Parameters ϕ derart, dass $c(t)$ an den Messzeitpunkten \bar{t}_i , $i = 1, \dots, k$, im Sinne der kleinsten Quadrate bestmöglich mit den diskreten Approximationen aus \hat{C} übereinstimmt.

C.1.1 Das chemische System

In Anhang B.2 erfolgte für den untersuchten Datensatz mittels PCD nur eine Bestimmung der Faktorenpaare der Katalysatorspezies Acylkomplex (folgend mit C bezeichnet) und Hydridokomplex (D). Zusätzlich umfasste das System noch das Edukt 3,3-Dimethyl-1-buten (B) und die Produkte 4,4-Dimethylpentanal und 2,3,3-Trimethylbutanal, welche für die folgenden Rechnungen zu einer Komponente (A) zusammengefasst werden. Weiter ist der Versuch zweimal unter gleichen Bedingungen durchgeführt worden, siehe Abbildung B.6. Es sei daran erinnert, dass von den Versuchen aufgrund von Mischungsvorgängen jeweils die Anfänge fehlen und somit $t \in [2.01 \text{ min}, 42.11 \text{ min}]$ beziehungsweise $t \in [2.01 \text{ min}, 40.20 \text{ min}]$ gilt.

C.1.2 Das kinetische Modell

Dem untersuchten Ausschnitt des Versuchs wird das Reaktionsmodell



zugrunde gelegt, was auf die Differentialgleichung

$$\begin{pmatrix} \dot{c}_A(t) \\ \dot{c}_B(t) \\ \dot{c}_C(t) \\ \dot{c}_D(t) \end{pmatrix} = \begin{pmatrix} & & & k_3 c_C(t) \\ -k_1 c_B(t) c_D(t) & +k_2 c_C(t) & & \\ k_1 c_B(t) c_D(t) & -k_2 c_C(t) & -k_3 c_C(t) & \\ -k_1 c_B(t) c_D(t) & +k_2 c_C(t) & +k_3 c_C(t) & \end{pmatrix} \quad (\text{C.2})$$

führt. Der Parameter ist $\phi = (k_1, k_2, k_3)^T$.

Die Anpassung der diskretisierten relativen Daten in \hat{C} auf absolute $C^{(\text{abs})}$ erfolgt mittels der Zusatzinformationen von zwei kalibrierten absoluten Faktoren $\hat{a}_C(x)$, $\hat{a}_B(x)$ und anschließender Skalierung von $\hat{c}_A(t)$ und $\hat{c}_D(t)$ über eine lineare Ausgleichsrechnung mittels Informationen über die Stoffumsätze.

C.1.3 Die Parametrisierung

Zur Bestimmung der Parameter werden die zwei in Abschnitt 7 erläuterten Methoden eingesetzt. Einerseits erfolgt die Anpassung über eine induktiv aufgebaute Folge von Anfangswertproblemen und andererseits über die Anpassung einer autonomen Folge von Anfangswertproblemen. Da zu dem System zwei Versuche unter gleichen Voraussetzungen durchgeführt wurden und damit aus zwei Faktorisierungen Daten für die anzupassenden Parameter vorliegen, wird jede Methode auf beide Datensätze angewendet.

Die Anpassung über eine induktive Folge von Anfangswertproblemen

Bei der Anwendung der Methode einer induktiv aufgebauten Folge von Anfangswertproblemen zur Parameterbestimmung ergeben sich als optimale Parameter

$$\phi^* = \begin{pmatrix} k_1 \\ k_2 \\ k_3 \end{pmatrix} = \begin{pmatrix} 1.6544 \text{ E} + 03 \text{ l mol}^{-1} \text{ min}^{-1} \\ 1.5808 \text{ E} - 04 \text{ min}^{-1} \\ 2.6837 \text{ E} + 02 \text{ min}^{-1} \end{pmatrix}$$

und

$$\tilde{\phi}^* = \begin{pmatrix} \tilde{k}_1 \\ \tilde{k}_2 \\ \tilde{k}_3 \end{pmatrix} = \begin{pmatrix} 1.4807 \text{ E} + 03 \text{ l mol}^{-1} \text{ min}^{-1} \\ 1.1344 \text{ E} - 06 \text{ min}^{-1} \\ 2.5877 \text{ E} + 02 \text{ min}^{-1} \end{pmatrix}.$$

Es bezieht sich ϕ^* auf den Versuch mit $k = 64$ Spektren und $\tilde{\phi}^*$ auf den mit $k = 61$ Spektren.

Die Anpassung über eine autonome Folge von Anfangswertproblemen

Als Vergleich dazu sind bei einer Systemanpassung mittels einer autonomen Folge von Anfangswertproblemen die optimalen Parametersätze

$$\phi^* = \begin{pmatrix} 1.4710 \text{ E} + 03 \text{ l mol}^{-1} \text{ min}^{-1} \\ 1.3506 \text{ E} - 00 \text{ min}^{-1} \\ 2.2127 \text{ E} + 02 \text{ min}^{-1} \end{pmatrix}, \quad \tilde{\phi}^* = \begin{pmatrix} 1.4594 \text{ E} + 03 \text{ l mol}^{-1} \text{ min}^{-1} \\ 8.6992 \text{ E} - 02 \text{ min}^{-1} \\ 2.1115 \text{ E} + 02 \text{ min}^{-1} \end{pmatrix}.$$

Die skalierten approximierten Teilfaktoren in \hat{C} aus Abschnitt B.2 sowie die Lösungen der angepassten Anfangswertprobleme sind in Abbildung C.1 graphisch dargestellt.

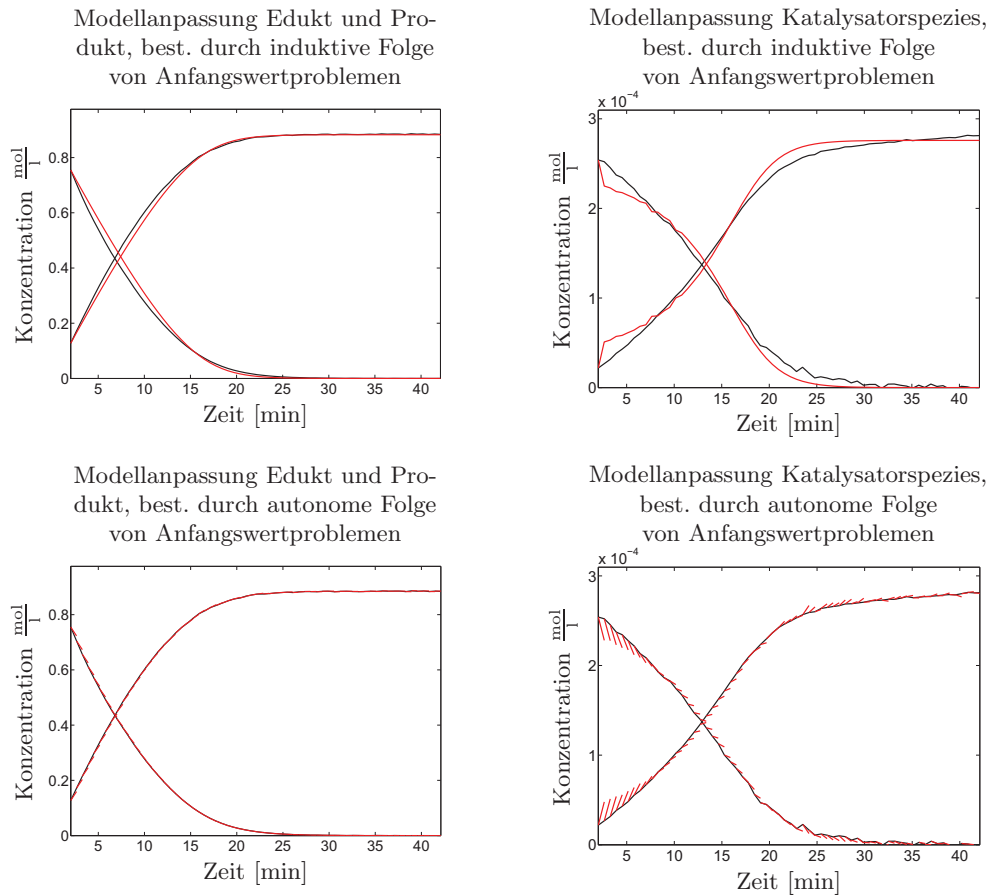


Abbildung C.1: Oben: Dargestellt ist die kinetische Anpassung mittels nichtlinearer Ausgleichsrechnung über eine induktive Folge von Anfangswertproblemen. Der kalibrierte Faktor $C^{(\text{abs})}$ ist schwarz, die numerische Lösung $C^{(\text{dgl})}$ des angepassten kinetischen Modells aus (C.2) ist rot dargestellt. Unten sind die Ergebnisse bei Anwendung der Methode einer autonomen Folge von Anfangswertproblemen abgebildet. Für beide Folgen ist nur die Anpassung des ersten Versuchablaufs dargestellt.

C.2 Die Anpassung eines kinetischen Modells zum System aus Anhang B.3

Dem Datensatz aus Abschnitt B.3 liegt ein System $A \rightarrow B \rightarrow C$ zugrunde. Bei der Messdatenaufnahme war das Edukt zu Reaktionsbeginn mit der Konzentration 0.013 mol l^{-1} vorhanden. Mit der Zusatzinformation des Stoffumsatzes $1-1-1$ lassen sich die approximierten Teilfaktoren \hat{C} zu $C^{(\text{abs})}$ skalieren.

C.2.1 Das kinetische Modell und die Parametrisierung

Für das vorliegende System ergibt sich das Differentialgleichungssystem

$$\begin{pmatrix} \dot{C}_A \\ \dot{C}_B \\ \dot{C}_C \end{pmatrix} = \begin{pmatrix} -k_1 C_A \\ k_1 C_A - k_2 C_B \\ k_2 C_B \end{pmatrix}.$$

Eine Anpassung dieses Modells an die kalibrierten Konzentrationsdaten erfolgt nur mit der

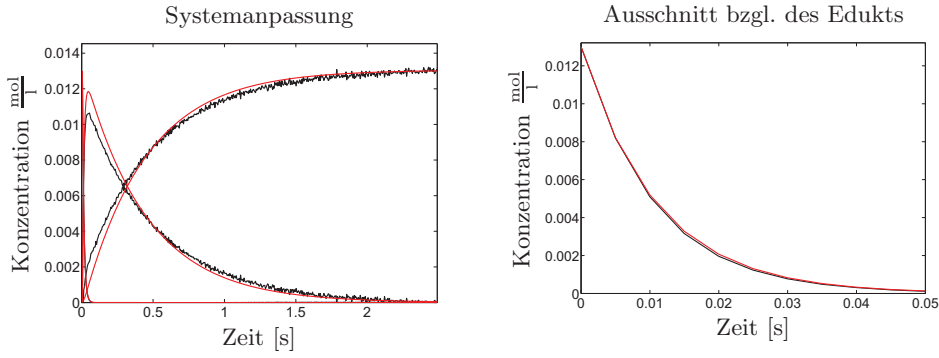


Abbildung C.2: Links: Dargestellt ist der kalibrierte Teilfaktor $C^{(\text{abs})}$ (schwarz) und die numerische Lösung des angepassten kinetischen Modells (rot). Rechts: Die Darstellung zeigt einen Ausschnitt des Edukts betreffend.

Methode einer induktiven Folge von Anfangswertproblemen und ergibt die Werte

$$\phi^* = \begin{pmatrix} k_1 \\ k_2 \end{pmatrix} = \begin{pmatrix} 91.929 \text{ s}^{-1} \\ 2.271 \text{ s}^{-1} \end{pmatrix}. \quad (\text{C.3})$$

Die kalibrierten Teilfaktoren $C^{(\text{abs})}$ sowie die numerische Lösung des sich aus dem parametrisierten Differentialgleichungssystem ergebenden Anfangswertproblems sind in Abbildung C.2 dargestellt.

C.2.2 Das Verhalten der Optimierungsfunktion in dem errechneten Parametersatz

Der Gradient der Optimierungsfunktion $G(\phi)$ aus (7.5) an dem berechneten Parametersatz ϕ^* sowie die Eigenwerte der Hessematrix

$$(\nabla G)(\phi^*) = \begin{pmatrix} 8.4622 \text{ E} - 06 \\ -1.2950 \text{ E} - 04 \end{pmatrix}, \quad \sigma(H_G(\phi^*)) = \begin{pmatrix} 1.9916 \text{ E} + 01 \\ 2.8869 \text{ E} - 01 \end{pmatrix},$$

sind im Rahmen einer numerischen Optimierung klare Kennzeichen einer lokalen Minimalstelle.

C.2.3 Analytische Lösung des kinetischen Modells

Die Lösung des, in diesem Abschnitt beschriebenen, kinetischen Modells lässt sich auch analytisch bestimmen und lautet ohne Berücksichtigung der Einheiten

$$c(t) = \frac{0.013k_2}{k_1 - k_2} \exp(-k_1 t) \begin{pmatrix} \frac{k_1 - k_2}{k_2} \\ -\frac{k_1}{k_2} \\ 1 \end{pmatrix} + \frac{0.013k_1}{k_2 - k_1} \exp(-k_2 t) \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} + 0.013 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Zur Bestimmung der Parameter k_1 und k_2 ist das nichtlineare Ausgleichsproblem

$$H(k_1, k_2) = \sum_{i=2}^k \sum_{j=1}^3 (C_{i,j}^{(\text{abs})} - c_j(\bar{t}_i))^2$$

zu lösen. Dabei sind die Einträge von $C^{(\text{abs})}$ die auf absolute Konzentrationen skalierten Werten aus der nichtnegativen Matrixfaktorisierung. Die optimalen Parameter zur Minimierung von H lauten

$$\begin{aligned} k_1 &= 94.403 \text{ s}^{-1}, \\ k_2 &= 2.2666 \text{ s}^{-1} \end{aligned}$$

und unterscheiden sich nur geringfügig von denen des numerischen Modells aus (C.3).

D Ergänzungen und Diagramme

D.1 Der Beweis von Korollar 4.23

Korollar 4.23. (Vergleiche ebenfalls Cohen, Rothblum, [13]) Aus Satz 4.21 lässt sich ableiten, dass jede nichtnegative Matrix $A \in \mathbb{R}_+^{k \times n}$ vom Rang zwei eine nichtnegative Matrixfaktorisierung besitzt. Insbesondere lässt sich eine solche Faktorisierung direkt berechnen.

Beweis: Der Beweis wird durch vollständige Induktion über die Anzahl der Spalten von A geführt. Es sei $A = (a_1, \dots, a_n)$, $a_i \in \mathbb{R}_+^k$, $i = 1, \dots, n$, derart, dass $\text{rank}((a_1, a_2)) = 2$ gilt (notfalls mittels Spaltenpermutation). Dass die Matrix $(a_1, a_2, a_3) \in \mathbb{R}_+^{k \times 3}$ eine nichtnegative Zerlegung $A = (a_1, a_2)Y$ besitzt, ist nach Satz 4.21 klar. Das ist der Induktionsanfang. Sei nun (möglicherweise wieder durch Anwendung einer Spaltenpermutation und mit anderem Y)

$$(a_1, \dots, a_z) = (a_1, a_2)Y.$$

Es reicht zu zeigen, dass es eine Faktorisierung von

$$A' = (a_1, \dots, a_{z+1}) = X'Y'$$

mittels $X' = (a_1, a_2)$ oder $X' = (a_{z+1}, a_2)$ und einem jeweils zugehörigen Y' gibt. (Der Fall $X' = (a_1, a_{z+1})$ ist so ebenfalls abgehandelt.) Es ist $\text{rank}(a_1, a_2, a_{z+1}) = 2$ und somit gibt es $\alpha \in \mathbb{R}^3 \setminus \{0\}$ mit $\alpha_1 a_1 + \alpha_2 a_2 + \alpha_3 a_{z+1} = 0$, wobei zwei Einträge in α größer oder gleich Null sind und einer kleiner Null ist. Es sind zwei Fälle zu betrachten. Der erste, $\alpha_1, \alpha_2 \geq 0, \alpha_3 < 0$, ist trivial. Dann gilt $a_{z+1} = -\frac{\alpha_1}{\alpha_3} a_1 - \frac{\alpha_2}{\alpha_3} a_2$ und es ist durch

$$A' = (a_1, a_2) \left(Y, \begin{pmatrix} -\frac{\alpha_1}{\alpha_3} \\ -\frac{\alpha_2}{\alpha_3} \end{pmatrix} \right)$$

eine nichtnegative Matrixfaktorisierung von A' gegeben. Der zweite Fall ist $\alpha_1 < 0, \alpha_2, \alpha_3 \geq 0$ (analog dazu wäre $\alpha_2 < 0, \alpha_1, \alpha_3 \geq 0$). Es gilt $a_1 = -\frac{\alpha_2}{\alpha_1} a_2 - \frac{\alpha_3}{\alpha_1} a_{z+1}$ und durch

$$A' = (a_{z+1}, a_2) \left(TY, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right)$$

mit der nichtnegativen Matrix

$$T = \begin{pmatrix} -\frac{\alpha_3}{\alpha_1} & 0 \\ -\frac{\alpha_2}{\alpha_1} & 1 \end{pmatrix}$$

ist eine nichtnegative Matrixfaktorisierung von A' gegeben. Der Rest ist durch Spaltenpermutationen lösbar. So ergibt sich auch direkt eine Faktorisierung der Ausgangsmatrix A . \square

D.2 Diagramme

Folgende Diagramme sollen die Abläufe in der Multi-Level-Strategie und bei der sukzessiven Faktorisierung sowie den Aufbau von PCD veranschaulichen.

D.2.1 Ablaufdiagramm zur Multi-Level-Strategie

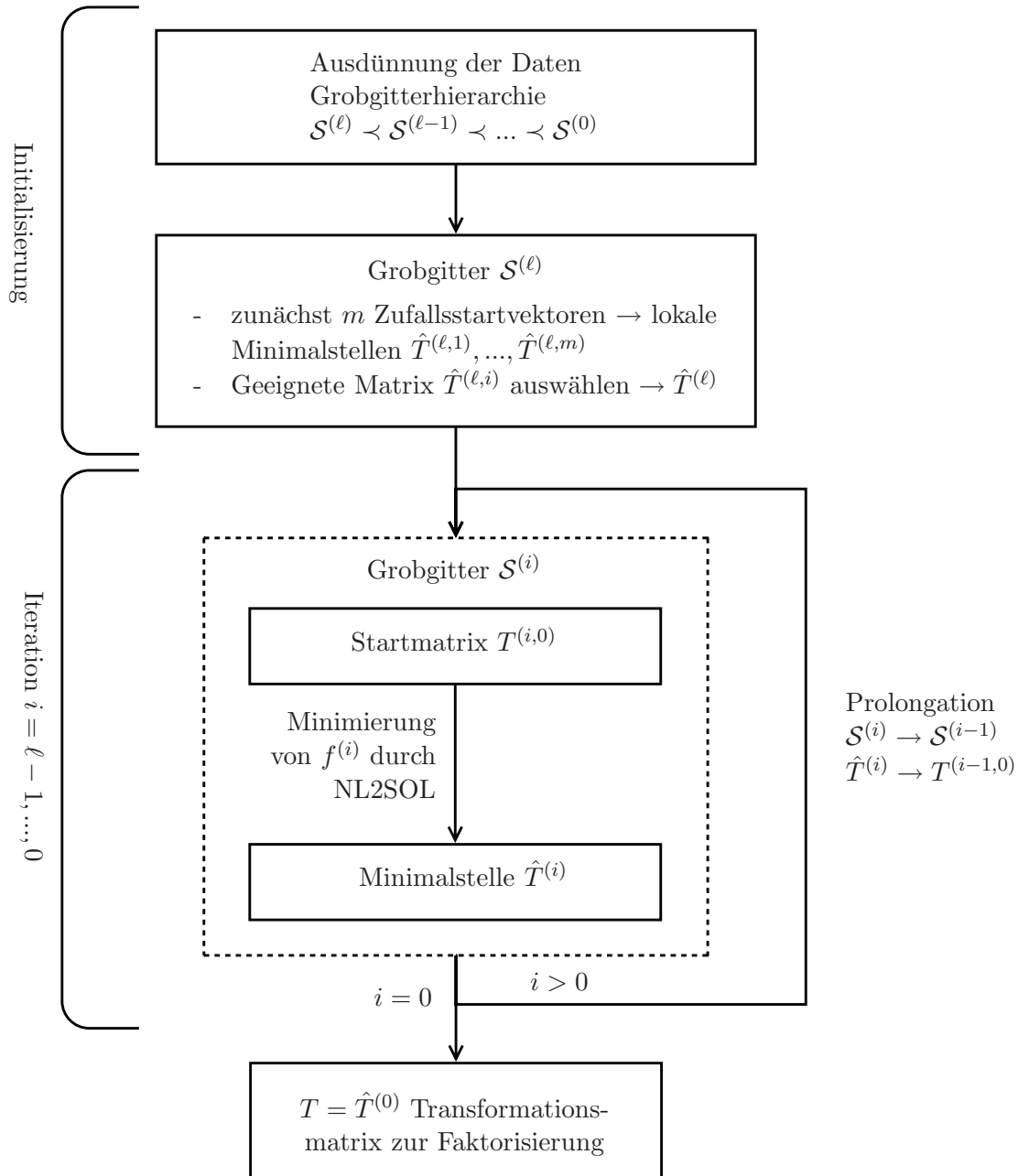


Abbildung D.1: Die Grafik zeigt das Ablaufdiagramm der Multi-Level-Strategie, beschrieben in Abschnitt 6.4.

D.2.2 Ablaufdiagramm zur sukzessiven Faktorisierung

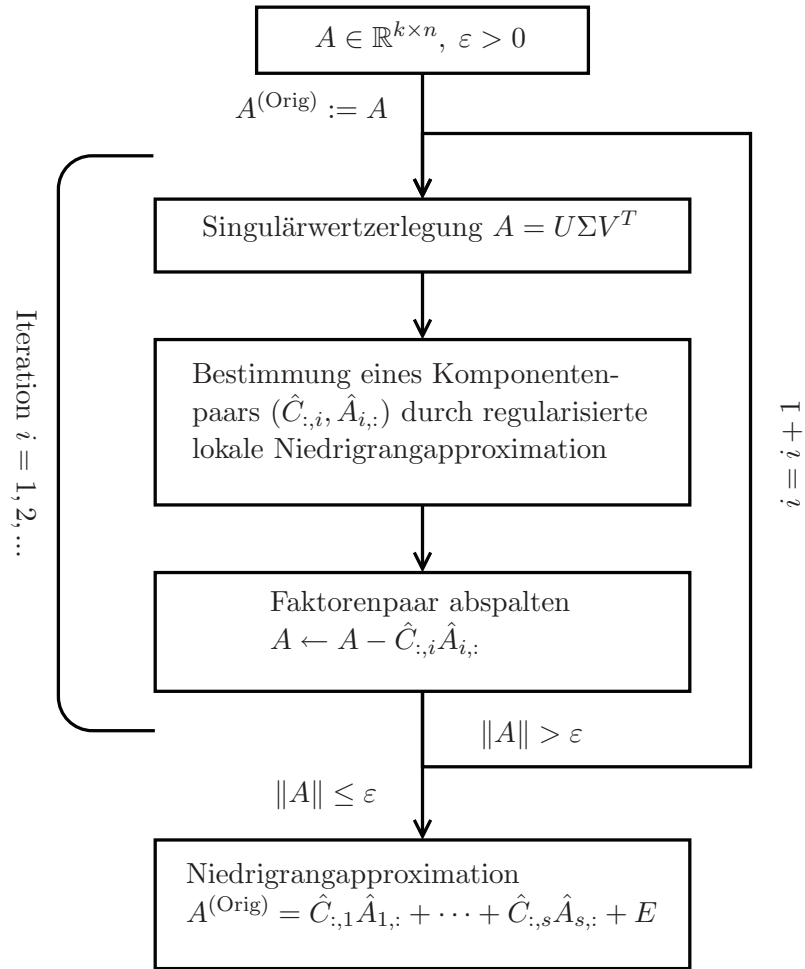


Abbildung D.2: Die Grafik zeigt das Ablaufdiagramm zur Matrixfaktorisierung durch einzelnes Abspalten der Faktoren, wie in Abschnitt 6.5 beschrieben.

D.2.3 Ablaufdiagramm zur Implementierung

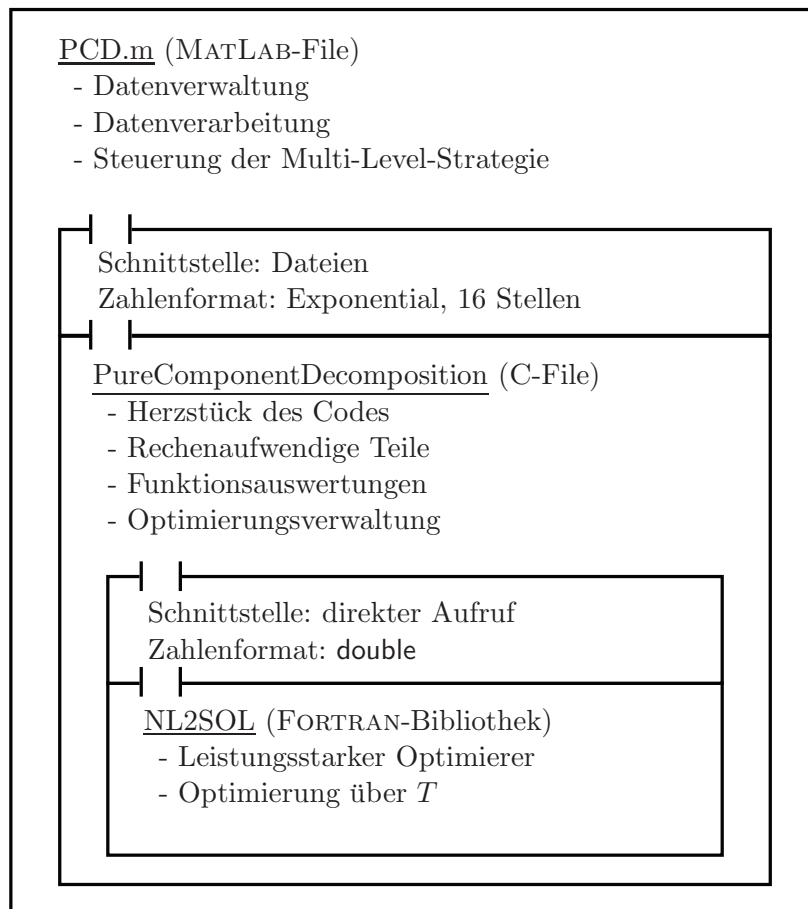


Abbildung D.3: Die Grafik zeigt den Programmaufbau. Übergeordnet ein MATLAB-File, darin integriert ein C-File, welches wiederum eine FORTRAN-Bibliothek enthält.

Literaturverzeichnis

- [1] U. M. Ascher, R. M. Mattheij, and R. D. Russell, *Numerical solutions of boundary value problems for ordinary differential equations*, SIAM, 1995.
- [2] M. W. Berry and M. Browne, *Email surveillance using non-negative matrix factorization*, Computational & Mathematical Organization Theory **11** (2005), 249–264.
- [3] M. W. Berry, M. Browne, A. N. Langville, V. P. Pauca, and R. J. Plemmons, *Algorithms and applications for approximate nonnegative matrix factorization*, Computational Statistics & Data Analysis **52** (2007), 155–173.
- [4] T. Beweries, C. Fischer, S. Peitz, V. V. Burlakov, P. Arndt, W. Baumann, A. Spannenberg, D. Heller, and U. Rosenthal, *Combination of spectroscopic methods: In situ NMR and UV/Vis measurements to understand the formation of group 4 metallacyclopentanes from the corresponding metallacyclopropenes*, Journal of the American Chemical Society **131** (2009), no. 12, 4463–4469.
- [5] Å. Björck, *Numerical methods for least squares problems*, SIAM, 1996.
- [6] R. G. Brereton, *Chemometrics: Data Analysis for the Laboratory and Chemical Plant*, Wiley, 2003.
- [7] D. Calvetti, S. Morigi, L. Reichel, and F. Sgallari, *Tikhonov regularization and the L-curve for large discrete ill-posed problems*, Journal of Computational and Applied Mathematics **123** (2000), no. 1-2, 423–446.
- [8] S. L. Campbell and G. D. Poole, *Computing nonnegative rank factorizations*, Linear Algebra and its Applications **35** (1981), 175–182.
- [9] C. Chatfield and A. J. Collins, *Introduction to multivariate analysis*, Chapman & Hall, 1996.
- [10] J.-C. Chen, *The nonnegative rank factorizations of nonnegative matrices*, Linear Algebra and its Applications **62** (1984), 207–217.
- [11] W. Chew, E. Widjaja, and M. Garland, *Band-target entropy minimization (BTEM): An advanced method for recovering unknown pure component spectra. Application to the FT-IR spectra of unstable organometallic mixtures*, American Chemical Society Organometallics **21** (2002), no. 9, 1982–1990.
- [12] A. Cichocki, R. Zdunek, and S. Amari, *New algorithms for non-negative matrix factorization in applications to blind source separation*, Proceedings of the IEEE international conference on acoustics, speech and signal processing, Toulouse, France (2006).
- [13] J. E. Cohen and U. G. Rothblum, *Nonnegative ranks, decompositions, and factorizations of nonnegative matrices*, Linear Algebra and its Applications **190** (1993), 149–168.
- [14] Pierre Comon, *Independent component analysis, A new concept?*, Signal Processing **36** (1994), no. 3, 287–314.
- [15] A. Corana, M. Marchesi, C. Martini, and S. Ridella, *Minimizing multimodal functions of continuous variables with the „simulated annealing“ algorithm*, ACM Transactions on Mathematical Software **13** (1987), no. 3, 262–280.
- [16] T. M. Cover and J. A. Thomas, *Elements of information theory*, Wiley, 1991.

- [17] J. E. Dennis, D. M. Gay, and R. E. Welsch, *NL2SOL—An adaptive nonlinear least squares algorithm*, ACM Transactions on Mathematical Software **7** (1981), no. 3, 369–383.
- [18] C. Eckard and G. Young, *The approximation of one matrix by another of lower rank*, Psychometrika **1** (1936), no. 3, 211–218.
- [19] H. W. Engl, M. Hanke, and A. Neubauer, *Regularization of inverse problems*, Kluwer Academic Publishers, 1996.
- [20] C. Fischer, T. Beweries, A. Preetz, H.-J. Drexler, W. Baumann, S. Peitz, U. Rosenthal, and Heller D., *Kinetic and mechanistic investigations in homogeneous catalysis using operando UV/Vis spectroscopy*, Catalysis Today **155** (2010), no. 3–4, 282–288.
- [21] O. Forster, *Analysis 2*, Vieweg, 1996.
- [22] G. H. Golub and C. F. van Loan, *Matrix computations, 3rd edition*, The Johns Hopkins University Press, 1996.
- [23] E. Gonzalez and Y. Zhang, *Accelerating the Lee–Seung algorithm for nonnegative matrix factorization*, Technical Report TR-05-02, Rice University (2005).
- [24] P. R. Griffiths and J. A. De Haseth, *Fourier transform infrared spectrometry, 2nd edition*, Wiley, 2007.
- [25] D. Guillaumet and J. Vitrià, *Determining a suitable metric when using non-negative matrix factorization*, 2002.
- [26] ———, *Non-negative matrix factorization for face recognition*, Lecture Notes in Computer Science **2504** (2002), 336–344.
- [27] ———, *Evaluation of distance metrics for recognition based on non-negative matrix factorization*, Pattern Recognition Letters **24** (2003), no. 9–10, 1599–1605.
- [28] H. Günzler and H. M. Heise, *IR-Spektroskopie, 3. Auflage*, VCH, 1996.
- [29] H. Günzler and A. Williams, *Handbook of Analytical Techniques, Volume 1*, Wiley-VCH, 2001.
- [30] W. Hackbusch, *Iterative Solution of Large Sparse Systems of Equations*, Springer, 1994.
- [31] E. Hairer and G. Wanner, *Solving ordinary differential equations II, 2nd edition*, Springer, 2002.
- [32] E. Hairer, G. Wanner, and S. P. Nørsett, *Solving ordinary differential equations I, 2nd edition*, Springer, 2002.
- [33] P. C. Hansen, *Truncated Singular Value Decomposition Solutions to Discrete Ill-Posed Problems with Ill-Determined Numerical Rank*, SIAM Journal on Scientific and Statistical Computing **11** (1990), no. 3, 503–518.
- [34] ———, *Analysis of Discrete Ill-Posed Problems by Means of the L-Curve*, SIAM Review **34** (1992), no. 4, 561–580.
- [35] P. C. Hansen and D. P. O’Leary, *The use of the L-curve in the regularization of discrete ill-posed problems*, SIAM Journal on Scientific and Statistical Computing **14** (1993), no. 6, 1487–1503.
- [36] B. Hofmann, *Mathematik inverser Probleme*, Teubner, 1999.
- [37] A. Hyvärinen, *Survey on independent component analysis*, Neural Computing Surveys **2** (1999), 94–128.
- [38] A. Hyvärinen, J. Karhunen, and E. Oja, *Independent component analysis: Algorithms and applications*, Neural Networks **13** (2000), 411–430.
- [39] ———, *Independent component analysis*, Wiley, 2001.

- [40] W. Härdle and L. Simar, *Applied multivariate statistical analysis*, Springer, 2007.
- [41] M. L. Jackson, *Soil chemical analysis, 2nd edition*, University of Wisconsin-Madison Libraries, 1985.
- [42] J. D. Jobson, *Applied multivariate data analysis, vol. 2: Categorical and multivariate methods*, Springer, 1992.
- [43] R. A. Johnson and D. W. Wichern, *Applied multivariate statistical analysis*, Pearson Education, 2002.
- [44] E. J. Karjalainen, *The spectrum reconstruction problem : Use of alternating regression for unexpected spectral components in two-dimensional spectroscopies*, Chemometrics and Intelligent Laboratory Systems **7** (1989), no. 1-2, 31–38.
- [45] J. Kauppinen and J. Partanen, *Fourier transforms in spectroscopy*, Wiley-VCH, 2000.
- [46] ———, *Fourier transforms in spectrometry*, Wiley-VCH, 2001.
- [47] H. Kim and H. Park, *Non-negative matrix factorization based on alternating non-negativity constrained least squares and active set method*, SIAM Journal on Matrix Analysis and Applications **30** (2008), 713–730.
- [48] H. Kim, H. Park, and L. Eldén, *Non-negative tensor factorization based on alternating large-scale non-negativity-constrained least squares*, Proceedings of the 7th IEEE international conference on bioinformatics & bioengineering (IEEE BIBE 2007) **2** (2007), 1147–1151.
- [49] P. Kosmol, *Methoden zur numerischen Behandlung nichtlinearer Gleichungen und Optimierungsaufgaben*, Teubner Studienbücher, 1989.
- [50] R. Kramer, *Chemometric techniques for quantitative analysis*, CRC Press, 1998.
- [51] J. B. Kruskal, *Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics*, Linear Algebra and its Applications **18** (1977), no. 2, 95–138.
- [52] W. J. Krzanowski and F. H. C. Marriott, *Multivariate analysis, part 1: Distributions, ordination and inference*, Wiley, 1994.
- [53] C. Kubis, R. Ludwig, M. Sawall, K. Neymeyr, A. Börner, K.-D. Wiese, D. Hess, R. Franke, and D. Selent, *A comparative in situ HP-FT-IR spectroscopic study of Bi- and Monodentate Phosphite-Modified Hydroformylation*, ChemCatChem **2** (2010), 287–295.
- [54] O. M. Kvalheim and Y.-Z. Liang, *Heuristic evolving latent projections: resolving two-way multicomponent data. 1. Selectivity, latent-projective graph, datascope, local rank, and unique resolution*, Analytical Chemistry **64** (1992), no. 8, 936–946.
- [55] C. L. Lawson and R. J. Hanson, *Solving least squares problems*, Prentice-Hall, Englewood Cliffs, 1974.
- [56] W. H. Lawton and E. A. Sylvestre, *Self modelling curve resolution*, Technometrics **13** (1971), 617–633.
- [57] D. D. Lee and H. S. Seung, *Learning the parts of objects by non-negative matrix factorization*, Nature **401** (1999), 788–791.
- [58] ———, *Algorithms for non-negative matrix factorization*, Advances in neural information processing systems **13** (2001), 556–562.
- [59] C.-J. Lin, *On the Convergence of Multiplicative Update Algorithms for Nonnegative Matrix Factorization*, Neural Networks, IEEE Transactions on **18** (2007), no. 6, 1589–1596.
- [60] ———, *Projected Gradient Methods for Nonnegative Matrix Factorization*, Neural Computation **19** (2007), no. 10, 2756–2779.

- [61] A. K. Louis, *Inverse und schlecht gestellte Probleme*, Teubner Studienbücher, 1989.
- [62] G. Maeß, *Vorlesungen über numerische Mathematik I*, Akademie Verlag, 1984.
- [63] E. R. Malinowski, *Factor analysis in chemistry, 3rd edition*, Wiley, 2002.
- [64] K. Neymeyr, M. Sawall, and D. Hess, *Pure component spectral recovery and constrained matrix factorizations: Concepts and applications*, Journal of Chemometrics **24** (2010), 67–74.
- [65] J. Nocedal and S. J. Wright, *Numerical optimization*, Springer, 1999.
- [66] P. Paatero, *A weighted non-negative least squares algorithm for three-way 'PARAFAC' factor analysis*, Chemometrics and Intelligent Laboratory Systems **38** (1997), no. 2, 223 – 242.
- [67] P. Paatero and U. Tapper, *Positive matrix factorization: A non-negative factor model with optimal utilization of error estimates of data values*, Environmetrics **5** (1994), 111–126.
- [68] Y. Pan, L. Susithra, and M. Garland, *Pure component reconstructions using entropy minimizations and variance-weighted piecewise-continuous spectral regions. Application to the unstable experimental system $\text{CO}_2(\text{CO})_8/\text{CO}_4(\text{CO})_{12}$* , Journal of Chemometrics **14** (2000), no. 2, 63–77.
- [69] A. Papoulis, *Probability, random variables and stochastic process, 3rd edition*, McGraw-Hill, 1991.
- [70] H. Park and H. Kim, *One-sided non-negative matrix factorization and non-negative centroid dimension reduction for text classification*, Proceedings of the workshop on text mining at the 6th SIAM international conference on data mining (2006).
- [71] V. P. Pauca, J. Piper, and R. J. Plemmons, *Nonnegative matrix factorization for spectral data analysis*, Linear Algebra and its Applications **416** (2006), no. 1, 29–47.
- [72] J. Polster, *Reaktionskinetische Auswertung spektroskopischer Meßdaten*, Vieweg, 1995.
- [73] V. Rakocëvić, *On continuity of the Moore-Penrose and Drazin inverses*, Mathematikum Bechnk **49** (1997), 163–172.
- [74] A. Rényi, *Wahrscheinlichkeitsrechnung mit einem Anhang über Informationstheorie, 5. Auflage*, VEB Deutscher Verlag der Wissenschaften, 1977.
- [75] P. Sajda, S. Du, T. Brown, and L. Parra, *Recovery of constituent spectra using non-negative matrix factorization*, Proceedings of SPIE **5207** (2003), 321–331.
- [76] C. E. Shannon, *A mathematical theory of communication*, ACM SIGMOBILE Mobile Computing and Communications Review **5** (2001), no. 1, 3–55.
- [77] A. Stegeman and N. D. Sidiropoulos, *On Kruskal's uniqueness condition for the Candecom/Parafac decomposition*, Linear Algebra and its Applications **420** (2007), no. 2-3, 540–552.
- [78] G. W. Stewart, *Perturbation theory for the singular value decomposition*. In: R. J. Vaccaro, editor, In: R. J. Vaccaro, Editor, [85]: SVD and signal processing II: Algorithms, analysis and applications (1991), 99–109.
- [79] ———, *Determining rank in the presence of error*, Tech. report, College Park, MD, USA, 1992.
- [80] ———, *On the early history of the singular value decomposition*, SIAM Review **35** (1993), 551–556.
- [81] J. Stoer and R. Burlisch, *Introduction to numerical analysis, 3rd edition*, Springer, 2002.
- [82] E. A. Sylvestre, W. H. Lawton, and M. S. Maggio, *Curve resolution using a postulated chemical reaction*, Technometrics **16** (1974), no. 3, 353–368.

- [83] B.-S. Tam, *A geometric treatment of generalized inverses and semigroups of nonnegative matrices*, Linear Algebra and its Applications **41** (1981), 225–272.
- [84] L. B. Thomas, *Rank factorization of nonnegative matrices (A. Berman)*, SIAM Review **16** (1974), no. 3, 393–394.
- [85] R. Vaccaro, *SVD and Signal Processing II: Algorithms, analysis and applications*, Elsevier, 1991.
- [86] E. Widjaja, C. Li, W. Chew, and M. Garland, *Band-target entropy minimization. A robust algorithm for pure component spectral recovery. Application to complex randomized mixtures of six components*, Analytical Chemistry **75** (2003), no. 17, 4499–4507.
- [87] R. Zdunek and A. Cichocki, *Non-negative Matrix Factorization with Quasi-Newton Optimization*, Artificial Intelligence and Soft Computing – ICAISC 2006 (L. Rutkowski, R. Ta-deusiewicz, L. Zadeh, and J. Zurada, eds.), Lecture Notes in Computer Science, vol. 4029, Springer Berlin / Heidelberg, 2006, pp. 870–879.
- [88] H. Zhang, W. Chew, and M. Garland, *Spectral pattern recognition of in situ FT-IR spectroscopic reaction data using minimization of entropy and spectral similarity (MESS): Application to the homogeneous rhodium catalyzed hydroformylation of isoprene*, Applied Spectroscopy **57** (2003), no. 5, 491–498.